



**HAL**  
open science

## **L'évaluation multicritères des systèmes d'élevage pour accompagner leurs évolutions : démarches, enjeux et questions soulevées**

Juliette Lairez, P. Feschet, Raphaëlle Botreau, Christian Bockstaller, Laurence Fortun-Lamothe, Isabelle Bouvarel, Joël Aubin

### ► To cite this version:

Juliette Lairez, P. Feschet, Raphaëlle Botreau, Christian Bockstaller, Laurence Fortun-Lamothe, et al.. L'évaluation multicritères des systèmes d'élevage pour accompagner leurs évolutions : démarches, enjeux et questions soulevées. INRA Productions Animales, 2017, 30 (3), pp.255-268. <hal-02617738>

**HAL Id: hal-02617738**

**<https://hal.inrae.fr/hal-02617738v1>**

Submitted on 25 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

# L'évaluation multicritère des systèmes d'élevage pour accompagner leurs évolutions : démarches, enjeux et questions soulevées

J. LAIREZ<sup>1</sup>, P. FESCHET<sup>2</sup>, R. BOTREAU<sup>3</sup>, C. BOCKSTALLER<sup>4</sup>,  
L. FORTUN-LAMOTHE<sup>5</sup>, I. BOUVAREL<sup>6</sup>, J. AUBIN<sup>7</sup>

<sup>1</sup> CIRAD, UR AIDA, TA B 115/02, Avenue Agropolis, 34398, Montpellier, France

<sup>2</sup> CIRAD, UPR GECCO, Campus agro-environnemental Caraïbe, 97285, Le Lamentin, Martinique

<sup>3</sup> Université Clermont Auvergne, INRA, Vetagro Sup, UMR Herbivores, 63122, Saint-Genès-Champagnelle, France

<sup>4</sup> UMR LAE, INRA, Université de Lorraine, 68000, Colmar, France

<sup>5</sup> GenPhySE, Université de Toulouse, INRA, INPT, INP-ENVT, 31320, Castanet Tolosan, France

<sup>6</sup> ITAVI, UMT, BIRD, 37380, Nouzilly, France

<sup>7</sup> SAS, Agrocampus Ouest, INRA, 35000, Rennes, France

Courriel : joel.aubin@inra.fr

La mise en place de démarches d'évaluation multicritère est une étape indispensable pour guider les systèmes d'élevage dans une dynamique de développement plus durable. Il existe aujourd'hui un foisonnement d'initiatives et d'outils d'évaluation ce qui rend difficile le choix d'une méthode pour l'évaluateur. Cet article relève un certain nombre de points clés pour la conception et la conduite d'évaluations multicritères des systèmes d'élevage.

L'élevage est un composant moteur de la durabilité de l'agriculture, du fait de ses externalités, positives et négatives sur les écosystèmes, les économies nationales et locales, et les sociétés. Il est reconnu pour fournir une alimentation de haute qualité nutritionnelle concernant les apports en protéines et en acides aminés, et est une source de revenu ou d'épargne. Par ailleurs, il contribue à maintenir l'emploi dans certaines régions rurales isolées et peut concourir à la reconnaissance d'un terroir à travers des appellations telles que les IGP et les AOP. Les productions animales jouent aussi un rôle majeur dans le fonctionnement des systèmes de productions végétales en apportant de la matière organique aux sols agricoles et en recyclant les nutriments de leurs effluents, réduisant ainsi leur dépendance aux intrants chimiques. En cela, l'élevage s'accorde à la définition d'une agriculture durable telle que définie par la FAO (2015), qui indique que l'agriculture durable doit favoriser des écosystèmes sains et une gestion durable de la terre, de l'eau et des ressources naturelles, tout en garantissant une sécurité alimentaire mondiale. L'agriculture durable doit également répondre aux besoins des générations présentes et futures quant aux produits et aux services, tout en garantissant une rentabilité, une

santé environnementale, et une équité sociale et économique. Comme pour l'ensemble des secteurs de productions, la déclinaison du concept de développement durable à l'agriculture a rapidement suscité de nombreuses interprétations et propositions (Hansen 1996, Zahm *et al* 2015). Un consensus s'est rapidement dégagé sur la nécessité de disposer de méthodes d'évaluation prenant en compte les différents enjeux du développement durable, donc de nature multicritère (encadré 1).

Malgré ces nombreux bénéfices, les productions animales sont aussi controversées car impliquées dans des problématiques environnementales, comme les marées vertes, le réchauffement climatique, ou encore les risques sur la santé humaine (zoonoses, antibio-résistance). L'augmentation de la consommation des produits animaux (viande, lait, œufs et produits transformés) au niveau mondial amplifie ces préoccupations. Tous ces éléments ont conduit à des exigences sociétales croissantes, que ce soit vis-à-vis du bien-être animal ou dans l'utilisation de modes d'élevage plus respectueux de l'environnement. Ceci est tout particulièrement vrai dans les pays occidentaux où la demande quantitative en produits animaux est satisfaite et où le niveau de

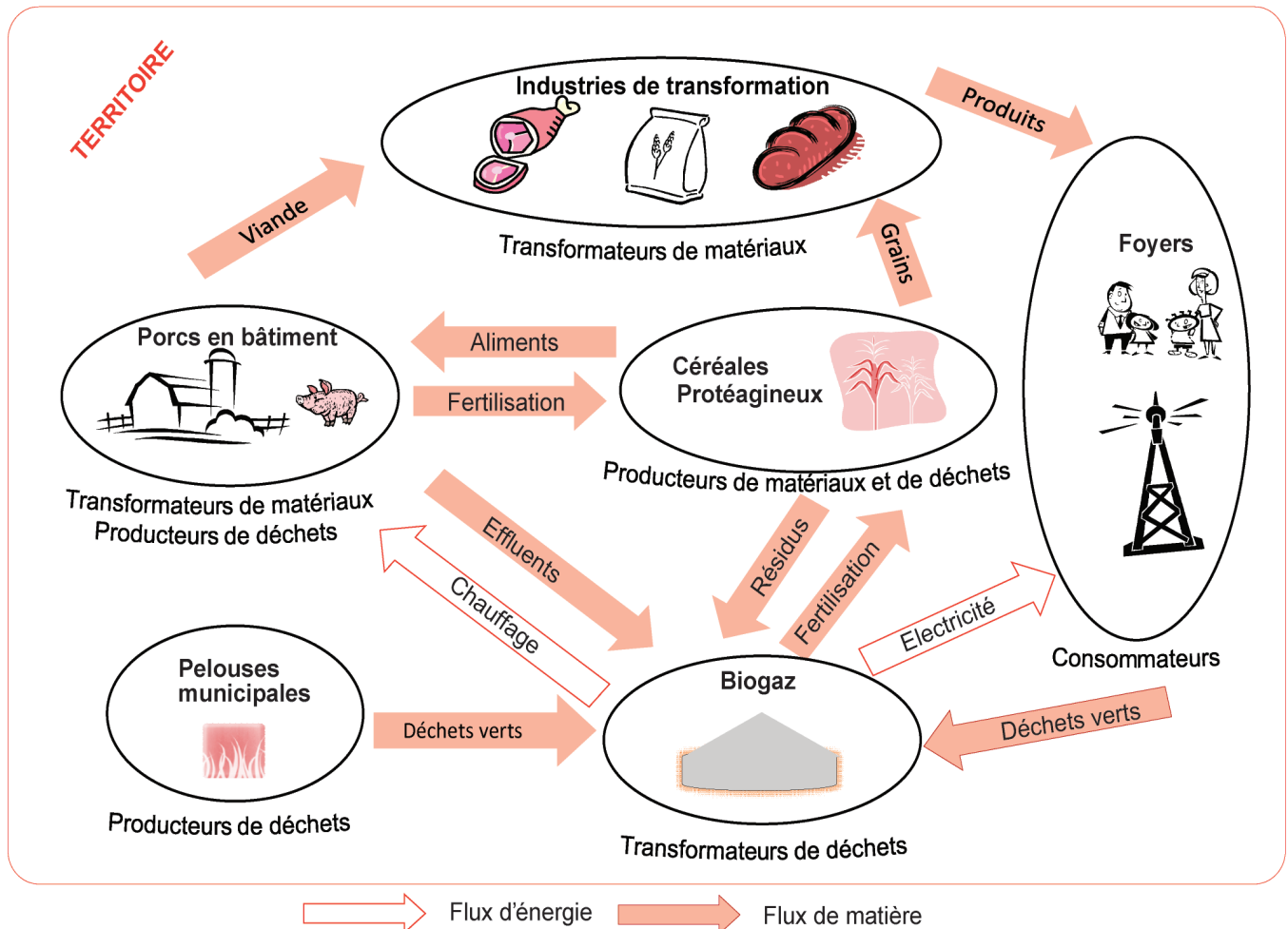
vie permet au consommateur de développer de nouvelles attentes. Ce contexte impose aujourd'hui à l'élevage d'évoluer vers des systèmes de production plus économes en intrants et plus en accord avec l'évolution du statut de l'animal, reconnu comme un être sensible, tout en assurant des performances économiques viables. Il s'agit, d'une part, de revoir le fonctionnement des ateliers d'élevages, en allant vers une écologisation des pratiques comme l'application des principes de l'agroécologie (Dumont *et al* 2012), et/ou d'autre part, de revoir son intégration territoriale, en considérant par exemple, l'atelier d'élevage comme un des maillons d'un réseau d'échanges de matières et de services entre activités humaines au sein d'un territoire (figure 1) comme le propose l'écologie industrielle et territoriale (Erkman 2004, Buclet 2011). Plus récemment, l'application des notions de services écosystémiques à l'élevage (Ryschawy *et al* 2015) ouvrent de nouvelles perspectives pour la définition des rôles multiples de l'élevage (encadré 2).

Cependant, toutes ces démarches nécessitent une clarification, notamment pour passer des enjeux et objectifs généraux énoncés à des critères déclinés en indicateurs permettant de faire un diagnostic. Tout l'enjeu de l'accompagnement

**Encadré 1.** L'évaluation multicritère.

Par opposition à l'évaluation monocritère, ne reposant que sur un seul critère (ex. la rentabilité économique), l'évaluation multicritère s'impose quand on cherche à comparer, hiérarchiser ou évaluer des options (par ex. des exploitations d'élevage) en prenant en compte plusieurs critères (ex. la rentabilité économique, la qualité de vie, la contribution au réchauffement climatique...). L'enjeu de l'évaluation multicritère est donc de fournir un jugement global sur la base d'une combinaison de différents éléments.

**Figure 1.** Exemple de l'écologie industrielle appliquée à l'élevage (adapté de Dumont et al 2012).

**Encadré 2.** Les services écosystémiques rendus par l'élevage.

L'évaluation des services rendus par l'élevage replace l'activité dans son contexte et l'associe à la perception des avantages (ou désavantages) pour l'homme, par les différents niveaux d'acteurs. Cette approche permet de considérer également les aspects positifs de l'élevage dans les contextes évalués. Le Millennium Ecosystem Assessment (2005) a défini les services rendus par les écosystèmes à l'Homme. Une application de ce concept à l'agriculture, et notamment à l'élevage, est donc une adaptation par le fait qu'elle ne concerne plus un écosystème naturel mais un système anthropisé. Dans le MEA, les services sont déclinés en 4 groupes distincts : Approvisionnement (s'appuyant sur un bénéfice-produit), Support (couvrant les fonctions d'habitat, de recyclage des nutriments, protection de faune et flore...), Régulation (régulations hydrologiques, biologiques, climat...) et Culturels (éducation, inspiration, attractivité, récréation...).

À partir de ce canevas, Ryschawy *et al* (2015) ont construit une approche participative pour quantifier et cartographier les services rendus par l'élevage en France métropolitaine. Les 4 groupes de services sont devenus : Vitalité territoriale, Approvisionnement, Qualité environnementale, Patrimoine et qualité de vie. Les services ont été déclinés en critères puis une sélection d'indicateurs a ensuite été réalisée afin de quantifier des bouquets de services associés à des territoires d'élevage (échelle du département). Ce travail a permis d'identifier des bouquets de services pour différentes zones d'élevage. Cette démarche s'est cependant heurtée à l'absence d'indicateurs quantifiables et à la disponibilité des informations pour renseigner les indicateurs. Au-delà de la spécificité du point de vue, celui des services, ce sont bien les mêmes défis méthodologiques que l'évaluation multicritère présentés dans cet article : système d'indicateurs, leur mesure, l'agrégation et la représentation des résultats.

par une démarche d'évaluation multicritère va reposer sur la capacité à produire un système d'évaluation cohérent, en adéquation avec les objectifs fixés par les porteurs d'enjeux et autres acteurs, et s'appuyant sur un ensemble de données mesurables. Une des difficultés de telles démarches d'évaluation est la disponibilité des outils de mesure et des données pour les renseigner.

En France, les systèmes de culture ont intégré depuis longtemps la dimension multicritère dans leurs évaluations (Bockstaller *et al* 2008). La prise en compte des rotations, avec les interactions qu'elles supposent dans le temps et l'espace, a notamment permis d'intégrer de nombreux critères pour évaluer la performance des systèmes de culture. Les travaux sur l'élevage ont abordé l'échelle du « système » plus tardivement (Landais et Balent 1993) et donc la question de l'évaluation multicritère également. L'élevage, comme objet d'évaluation multicritère, représente pourtant un niveau d'intégration systémique élevé, notamment parce qu'il intègre les productions végétales pour l'alimentation animale et qu'il joue un rôle central dans le fonctionnement des cycles (carbone, azote, phosphore, eau...). Pour toutes ces raisons, l'évaluation multicritère des systèmes d'élevage, représente un enjeu à la fois pour les scientifiques, les acteurs du développement et la société. En effet, différents types d'acteurs, agriculteurs, conseillers, chercheurs, gestionnaires des territoires, citoyens et politiques sont en demande d'outils d'évaluation pour comprendre et décider. Ceci a conduit à un foisonnement d'initiatives comme des listes d'indicateurs ou des méthodes plus structurées et l'utilisateur est souvent démuné pour faire son choix (Bockstaller *et al* 2009, Schader *et al* 2014). Par ailleurs malgré cette multiplicité d'initiatives, des questions fondamentales subsistent encore au niveau scientifique : Comment évaluer les lacunes et points forts et la dynamique d'un système au regard de la durabilité ? Pour aller plus loin, comment traduire ces mesures en une information synthétique ? Comment, à partir de ces informations, prendre des décisions pour faire des choix opérationnels, puis faire évoluer les systèmes d'élevage ? En outre, il est parfois difficile de comprendre le domaine de validité et la robustesse des conclusions. Ces constatations ont conduit à l'édition d'un guide sur l'évaluation multicritère des systèmes de grandes cultures et d'élevage (Lairez *et al* 2015). À la suite de cet ouvrage, il nous a paru nécessaire de présenter les points méthodologiques clefs liés à l'évaluation multicritère et d'aborder les questions spécifiques de son application aux systèmes d'élevage pour l'accompagnement de leur transition.

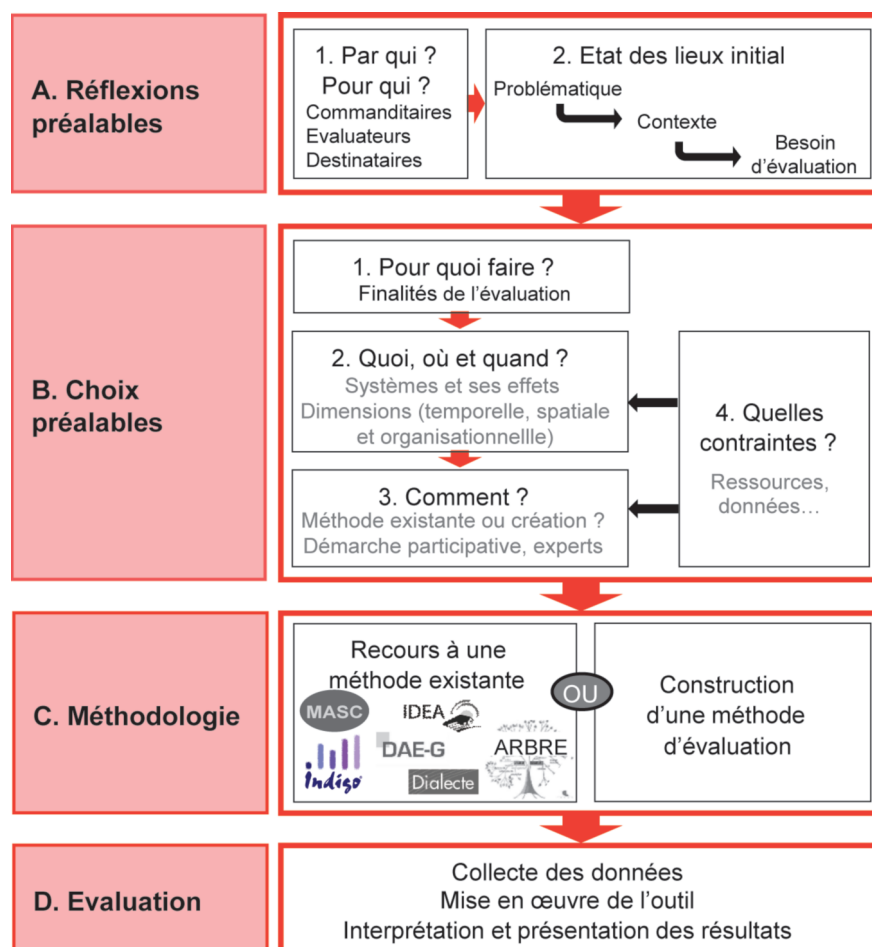
## 1 / Mettre de la rigueur dans la démarche d'évaluation

L'évaluation multicritère est un domaine en pleine effervescence. Les outils et les méthodes d'évaluation se développent abondamment, se différenciant sur le plan des enjeux traités et des critères évalués, des systèmes étudiés, des acteurs impliqués ou encore de la représentation des résultats. Certaines méthodes s'affichent comme étant très génériques (ex. IDEA, Vilain *et al* 2008) alors que d'autres revendiquent un développement par et pour des acteurs dans un contexte très localisé (ex. Laurent *et al* 2017, avec une co-construction d'une grille d'évaluation appliquée aux élevages laitiers en zone AOP Cantal). Avant de s'engager dans une stratégie d'évaluation, il convient d'adopter une démarche d'évaluation claire et transparente plutôt que de décider immédiatement d'utiliser une méthode existante ou d'en construire une nouvelle (Bockstaller *et al* 2013).

Cette démarche se déroule en quatre étapes (figure 2). La première étape

(étape A) consiste à identifier les acteurs de l'évaluation et circonscrire la problématique et le besoin d'évaluation. La deuxième étape (étape B) est une étape de choix puisqu'elle doit déterminer les objectifs et les attentes vis à vis de cette évaluation, c'est-à-dire les finalités, ainsi que l'utilisation qui sera faite des résultats. Dès lors, il est possible d'identifier l'objet et le périmètre de l'évaluation. Enfin, il faut déterminer la stratégie d'évaluation (recours à une méthode existante ou développement d'une nouvelle), ceci étant conditionné par des contraintes de ressources (financières, humaines, temporelles) et de données (disponibilité, accessibilité). La troisième étape (étape C) porte sur la définition du cadre conceptuel et méthodologique qui structure les critères et les indicateurs, en lien avec la méthode utilisée ou créée. Enfin la dernière étape (étape D) désigne la phase d'évaluation proprement dite. Le contenu de ces étapes a été présenté en détail par Lairez *et al* (2015). Nous proposons de revenir ci-après sur certains éléments fondamentaux de cette démarche.

Figure 2. Démarche générale d'évaluation d'après Lairez *et al* (2015).



La démarche est ici présentée de manière très linéaire, mais dans les faits, en avançant dans la démarche, le concepteur peut être amené à revenir sur des choix faits à des étapes préalables.

## 1.1 / Identification des finalités de l'évaluation multicritère

Il existe trois principales finalités génériques à l'évaluation, qu'elle soit multicritère ou non : connaître/s'informer, aider à la décision et communiquer ; ces trois finalités pouvant se combiner. Suivant les finalités de l'évaluation, différentes démarches peuvent être adoptées, ce qui détermine les outils à mobiliser et le niveau de précision des informations à agréger. En fonction du contexte une évaluation pourra être descriptive et informer sur un état ou sur une trajectoire, normative et informer sur l'adéquation par rapport à un référentiel, ou encore prescriptive avec un objectif de recommandation ou d'injonction sous-jacente. Les trois finalités sont illustrées ci-dessous.

### a) L'évaluation multicritère pour connaître les systèmes et dresser un diagnostic

Pour les systèmes agricoles, le caractère multicritère d'une évaluation est devenu incontournable. Il s'agit certes, d'aborder les différents piliers du développement durable, mais aussi d'apporter une vision structurée des caractéristiques du système. Cette évaluation descriptive peut alors être appliquée directement sur des systèmes existants et permettre un diagnostic. Le choix des indicateurs doit alors permettre de rendre compte d'un état. Aujourd'hui la plupart des évaluations multicritères sont synchroniques, avec par exemple des données collectées sur une année comptable agricole pour la comparaison de plusieurs exploitations à un instant « t » (Lairez *et al* 2015). Ces évaluations définissent des seuils relatifs, il peut s'agir par exemple des meilleures valeurs prises pour chaque indicateur par un ensemble d'exploitations d'une région agricole. Le système le plus durable sera celui qui se rapproche le plus de ces valeurs. Il est plus rare d'évaluer la durabilité des trajectoires d'évolution des systèmes d'élevage, car celle-ci nécessite des données sur plusieurs années ou fait appel à la modélisation. Le développement durable étant un processus continu, la durabilité des systèmes d'élevage doit s'appréhender plutôt à travers l'analyse de leurs trajectoires d'évolution. Ces trajectoires peuvent s'infléchir sous la pression de différentes contraintes comme le changement climatique (Dong *et al* 2015), la variation des prix agricoles, ou encore des choix politiques. Au travers d'une analyse multicritère, il est possible de mettre en évidence des modèles d'évolution qui allient un ensemble de caractéristiques des systèmes, comme le montrent (García-Martínez *et al* 2008) pour les élevages de montagne en Europe, ou pour les systèmes de polyculture-élevage des coteaux de Gascogne (Ryschawy *et*

*al* 2012). La difficulté de ce type d'évaluations prospectives est de prendre en compte les effets de changement de contexte de la production sur l'évolution des systèmes d'élevage. Certaines études évaluent les effets de l'évolution du contexte extérieur sur différents indicateurs de durabilité des exploitations (Belhouchette *et al* 2011, Viaggi *et al* 2013, Reidsma *et al* 2015, Wolf *et al* 2015). Belhouchette *et al* (2011) ont par exemple utilisé un modèle bioéconomique pour évaluer les effets de la directive Nitrates sur trois types d'exploitation. Les indicateurs évalués étaient le revenu agricole, la lixiviation des nitrates, l'érosion et la consommation d'eau. Viaggi *et al* (2013) ont quant à eux évalué les effets du découplage des aides de la PAC sur 3 indicateurs (le revenu agricole, l'emploi et l'utilisation des fertilisants chimiques à l'échelle des exploitations).

D'autres échelles d'évolution nécessitent aussi un accompagnement par des évaluations multicritères. À l'échelle des territoires et des filières, les évaluations multicritères sont utilisées dans le cadre de diagnostics et peuvent aussi renseigner sur les attentes des parties prenantes. La méthode OVALI permet ainsi d'exprimer les attentes de la filière avicole et de la société et d'identifier des pistes d'amélioration. Par exemple, sur la base d'un cas-type de la filière Poulet standard en Pays de la Loire, un scénario combinant rénovation et construction des bâtiments, évolutions des caractéristiques nutritionnelles des aliments et de la gestion du lot et l'approvisionnement local en matières premières indique une possible amélioration des coûts de production et des performances environnementales, et une amélioration de l'image du produit (Dusart *et al* 2015).

Dans la pratique ces évaluations posent un grand nombre de problèmes méthodologiques, notamment lors d'évaluations en continu qui interrogent sur les valeurs de référence permettant d'interpréter les indicateurs. Dans tous les cas, l'évaluation descriptive constitue la base nécessaire à toute évaluation multicritère et peut être complétée pour aller vers une évaluation normative voire prescriptive. Une évaluation normative vise à produire un résultat synthétique se basant sur la comparaison à un référentiel (norme) permettant de savoir si le résultat est bon ou mauvais en termes de durabilité. Ce type d'évaluation est utilisé lorsque l'on veut produire une note de durabilité, un classement sur une échelle de notation ou encore une certification pour un label. Cette évaluation peut se faire à l'échelle d'une exploitation mais aussi à des échelles plus vastes, comme par exemple celle de la filière quand l'évaluation porte sur le produit vendu au consommateur.

### b) L'évaluation multicritère pour aider à la décision et accompagner l'innovation et la transition

L'enjeu de l'accompagnement de la transition des systèmes d'élevage, est de maintenir ou d'augmenter leur durabilité dans le cadre de changements de contexte (changement climatique, fluctuation des prix, nouvelles politiques) ou d'évolution d'objectifs tactiques et stratégiques de l'agriculteur (par exemple autonomie fourragère, introduction d'un robot de traite). Dans ce cas, l'évaluation multicritère est utilisée en vue d'une adaptation/évolution des systèmes d'élevage à ces changements. Pour accompagner l'innovation dans les systèmes de production, deux niveaux sont à différencier selon que l'innovation porte sur un changement technique isolé ou un changement global du système évalué.

Dans le cas d'évolutions techniques, l'évaluation dressera un bilan des effets positifs et négatifs ainsi que des avantages et contraintes des systèmes actuels afin de définir des objectifs opérationnels pour la mise en œuvre du changement de technique. Ce peut être le cas dans un changement dans le mode de distribution des aliments par exemple.

Certaines évolutions techniques ont pour conséquences la reconfiguration de l'ensemble du système de production. C'est le cas pour les ruptures technologiques comme l'installation d'un robot de traite qui peut conduire à revoir la conduite de l'alimentation, le parcellaire ou la génétique du troupeau. Pour accompagner l'évolution du système de production, il faudra prendre en compte dans l'évaluation la reconfiguration de l'ensemble des techniques, mettre en évidence les synergies et compromis et évaluer les gains et risques potentiels liés à ces évolutions de pratiques. C'est le cas des transferts d'impacts souvent abordés en Analyse du Cycle de Vie (ACV), qui montrent par exemple que si le traitement des lisiers a un effet bénéfique sur la qualité de l'eau, il peut induire une augmentation des émissions de gaz à effet de serre. Cette évaluation à l'échelle du système se conduit aussi lors de changements globaux du système de production, comme ceux associés à un changement d'objectif de l'élevage par exemple dans le cas du passage d'un système de production dit conventionnel au cahier des charges de l'agriculture biologique ou d'un système de type « maïs-concentré » à un système dit « tout à l'herbe » en production laitière.

Pour accompagner la transition, une première approche est celle « pas à pas » de diagnostic/action, où chaque année des indicateurs/critères de performances et impacts du système d'élevage sont

évalués et des changements techniques ou plus globaux sont mis en place sur l'exploitation pour améliorer progressivement les résultats de l'évaluation d'année en année. Une seconde approche est d'imaginer des systèmes d'élevage innovants, en rupture avec les systèmes actuels. La pertinence de ces innovations est évaluée de manière *ex ante* à l'aide d'indicateurs. Giller *et al* (2011) ont développé un cadre d'analyse participatif pour la mise en place de boucles de co-innovation : l'approche DEED (« *describe, explain, explore and design* »). Elle consiste en une première étape de description du contexte (climat, sol, économie) et des systèmes de production (typologie). Une seconde étape d'explication pour analyser les problèmes existants. Ensuite, ils explorent différents scénarios de changement, puis ils mettent en œuvre les alternatives en réalisant un suivi-évaluation chez les agriculteurs. Un autre exemple d'approche pour accompagner les transitions de l'élevage est le « RAMI fourrager » développé par l'INRA. Il s'agit d'un jeu de plateau à faire avec un groupe d'éleveurs pour la conception de systèmes fourragers pour atteindre de nouveaux objectifs (se convertir au biologique, autonomie fourragère, nouveau cahier des charges à respecter) ou pour s'adapter à des changements de contexte (changement climatique, sécheresse) (Martin 2015). Un ensemble d'indicateurs de performance est fourni pour chaque changement afin d'accompagner le processus décisionnel des agriculteurs.

Pour redéfinir de nouvelles modalités d'intégration culture-pâture-élevage à l'échelle des territoires, Moraine *et al* (2014) proposent d'utiliser une méthode de co-construction avec les acteurs en trois étapes : diagnostic (définition des enjeux), brainstorming (identification de pistes) et évaluation (analyse multicritère). La troisième étape permet de mettre en évidence les points de blocage à la mise en place de nouvelles pratiques et d'identifier les principaux compromis à trancher et verrouillages à lever. L'évaluation multicritère est donc considérée ici comme une des étapes à part entière de la conception participative. L'intégration culture-élevage peut en effet prendre différentes formes d'intensité et d'étendue spatiale croissantes : de la simple complémentarité entre voisins à la synergie territoriale. Lorsque l'on se place à l'échelle du territoire, les évolutions engagent une grande diversité d'acteurs aux intérêts, mode de fonctionnement et/ou dynamique divergents (producteurs, coopératives, conseillers, gestionnaires...). Les méthodes de conception participative intégrant une phase d'évaluation multicritère permettent aux acteurs impliqués de définir des objectifs communs, de proposer des in-

dicateurs adaptés localement et de débattre des valeurs et coefficients à attribuer à chaque indicateur. L'évaluation multicritère permet ici d'engager l'action. Pour les bénéficiaires, il s'agit d'innover pour maximiser les intérêts de l'intégration culture-élevage. Pour les instances de gouvernance, il s'agit d'accompagner la transition en pensant l'intégration culture-élevage dans un cadre socio-écologique et en définissant des indicateurs de suivi de ces nouvelles formes de coordination territoriale. Par exemple, dans un territoire du sud-ouest de la France, un diagnostic initial a fait apparaître que, dans la zone amont (Ségala), les systèmes d'élevage sont très dépendants des ressources alimentaires externes alors que, dans la zone aval (plaine de Montauban), les systèmes de culture (maïs, blé/tournesol) sont très dépendants de la fertilisation azotée et sensibles au déficit hydrique. Les acteurs impliqués (agriculteurs, conseillers de chambres d'agriculture, de coopératives, responsables d'agences de l'eau) ont conçu un scénario de mise en place d'une filière de luzerne dans la zone aval qui permet de limiter l'utilisation des engrais azotés, de réguler les bioagresseurs, d'améliorer la fertilité des sols, et de réduire les besoins en eau. Elle permet en plus aux systèmes d'élevage de la zone amont d'être moins dépendants des approvisionnements étrangers et ainsi d'améliorer leur bilan carbone.

### c) L'évaluation multicritère pour communiquer

C'est le troisième objectif générique de l'évaluation. Il est souvent associé à l'un des deux précédents. Il est possible de différencier deux grands types d'évaluation associés à la communication, ceux qui s'appuient sur une promotion sans but directement marchand et qui permettent de promouvoir une image ou des valeurs (ex : campagne télévisuelle sur la viande de bœuf) ; et ceux qui ont un objectif de positionnement sur un marché et qui vont s'appuyer sur des normes, des chartes ou des démarches de certification. Les organisations des filières ont par exemple recours aux évaluations multicritères pour évaluer leurs performances (économiques, sociales, environnementales) et mettre en évidence les maillons à renforcer. Ces éléments peuvent servir à la communi-

cation en interne ou vers des tiers et conduire à des labellisations. Le projet AgriBalyse de l'ADEME d'inventaire du cycle de vie des produits agricoles français visant à l'affichage environnemental entre dans ce cadre (Koch et Salou 2014). Grâce à cet affichage, les consommateurs pourraient, s'ils le souhaitent, choisir lors de leurs achats des produits plus respectueux de l'environnement et ainsi permettre de favoriser les systèmes d'élevage dont ces produits sont issus. Ainsi, l'évaluation multicritère peut être utilisée pour orienter ces transformations et pour accompagner la transition de l'élevage vers des systèmes plus durables en s'appuyant sur des démarches de communication auprès des consommateurs.

## 1.2 / Définition du périmètre de l'évaluation

La détermination du périmètre de l'évaluation est délicate (cf. encadré 3) car elle implique de définir plusieurs éléments (frontières du système, flux entrants, flux sortants, effets...). Dans la plupart des méthodes d'évaluation multicritère, ces éléments ne font pas l'objet d'une réflexion particulière. Ils sont parfois partiellement explicites (ex : IDEA, Vilain *et al* 2008), mais le plus souvent le système étudié est réduit au « processus central de production » sans prendre en compte les processus impliqués en amont et en aval de la production (ex. matières premières, transport, utilisation des produits). Par exemple, l'évaluation se focalisera sur l'atelier bovin d'une exploitation laitière, sans prendre en compte les émissions de gaz à effet de serre issues du transport du lait jusqu'à la coopérative. Dans certaines méthodes comme IDEA ou INDIGO, où les indicateurs sont prédéfinis par la méthode, les processus en amont ne sont pris en compte que pour la consommation énergétique (Bockstaller *et al* 2013). Dans les méthodes plus génériques, comme l'Analyse du ACV (van der Werf et Petit 2002), la phase de définition du périmètre de l'évaluation constitue la première étape.

Le système étudié se définit à partir d'un ensemble de processus (e.g. alimentation, gestion des effluents) qui nécessitent des flux entrants (e.g. ressources

### Encadré 3. De la complexité de la définition du périmètre de l'évaluation.

Il existe aujourd'hui de nombreuses questions qui restent en suspens concernant la définition du périmètre de l'évaluation : comment faire pour déterminer le périmètre de l'évaluation lorsque les processus à considérer sont pertinents pour une dimension du développement durable mais pas pour une autre ? Comment choisir un niveau d'évaluation (exploitation, parcelle, atelier) pertinent pour tous les effets du système ? Comment faire pour déterminer le périmètre de l'évaluation lorsque les effets considérés se situent à de multiples niveaux ?

naturelles, infrastructures, intrants) et qui génèrent des flux sortants (e.g. émissions gazeuses, polluants, revenus). Ces flux engendrent des effets, c'est-à-dire des perturbations et modifications d'un milieu naturel ou socio-économique. Le système étudié est lui-même composé de sous-systèmes en interaction et s'inscrit dans des systèmes plus larges, un système local (e.g. exploitation, bassin versant, région) et un système global (ex. continent, planète). Ainsi il est source d'effets qui l'affectent lui-même tout en contribuant également aux enjeux des systèmes locaux (e.g. contribution à la qualité de l'eau via l'émission de nitrate) et globaux (e.g. contribution au réchauffement climatique mondial via l'émission de méthane). Le système et les effets se caractérisent par une dimension spatiale (délimitation physique), temporelle (période d'analyse) et organisationnelle (acteurs concernés).

La réalité étant par nature impossible à appréhender dans toute sa complexité, il n'est pas envisageable de rendre compte intégralement de l'ensemble du système et de ses effets. Le système évalué est donc une représentation la plus objective possible de cette réalité au regard des finalités étudiées. Dans l'élaboration d'une démarche d'évaluation, il est donc nécessaire de bien identifier au préalable les finalités de l'évaluation car selon les besoins, les échelles spatio-temporelles, définies par le grain (niveau de détail de l'information) et l'étendue de l'évaluation (limites spatiale ou temporelle étudiées, et nombre d'effets mesurés), seront plus ou moins fines ou larges. La détermination du périmètre de l'évaluation consiste ensuite à réaliser deux étapes :

#### a) *Étape 1 : Le choix des frontières du système évalué*

Il s'agit ici de définir les frontières c'est-à-dire les processus évalués et leurs flux entrants, ainsi que les dimensions spatiale, temporelle et organisationnelle du système. Cette étape est souvent délicate étant donné que les processus pertinents pour une évaluation sociale ne le sont pas forcément pour une évaluation environnementale. Par exemple, le désaisonnement d'un lot de brebis est en partie bénéfique car il permet d'étaler les mises bas et donc de mieux valoriser la ressource en herbe. Néanmoins, la gestion de l'alimentation de deux lots ayant à un instant des besoins différents et le dédoublement des périodes de pic d'activité, notamment au moment des agnelages, influencent la charge de travail de l'éleveur. Aussi, le choix de la dimension spatiale du système se fait en fonction des finalités de l'évaluation. Par exemple, il est possible, mais pas exclusif, de rendre compte des effets d'une politique agricole

à l'échelle d'un territoire alors que la comparaison de systèmes de production dans le cadre d'un diagnostic se fera à l'échelle d'une exploitation ou de l'atelier d'élevage. La dimension temporelle est également délicate à déterminer à cause de la validité des données et des indicateurs dans le temps puisque l'agriculture est en constante évolution (marchés, climat, organisation des acteurs...). Il convient de s'assurer de la validité des données utilisées, c'est-à-dire de leur représentativité (vis-à-vis du système, du temps et de l'espace), leur précision, et leur « date de péremption ».

#### b) *Étape 2 : Le choix des flux sortants du système et des effets à évaluer*

Il s'agit ici de prendre en compte les changements d'état à l'extérieur et au sein du système, dans les différentes dimensions spatiale, temporelle et organisationnelle du système. En ACV environnementale, des règles de décision, dénommées « critères de coupure », permettent d'éviter de choisir arbitrairement les flux et les effets à évaluer. Par exemple, si chacun des flux sortants ne dépasse pas 1% du total des impacts et que leur somme n'excède pas 5%, ces derniers ne sont généralement pas comptabilisés. Le choix des flux sortants et des effets à évaluer se fait souvent *a priori*, en s'appuyant sur des hypothèses établies à partir d'autres situations similaires préalablement évaluées. Le niveau pertinent pour évaluer à la fois les effets sociaux, économiques ou environnementaux n'est pas obligatoirement le même pour ces trois dimensions. Les effets évalués peuvent s'observer à un niveau supérieur à celui auquel les données sont collectées. C'est le cas par exemple d'une évaluation des effets d'un atelier d'élevage sur le changement climatique, sur la perte de biodiversité, ou encore sur la revitalisation des espaces ruraux. Une solution est d'évaluer la « contribution » des niveaux inférieurs (ici, l'atelier d'élevage) à ces effets plus larges. Il faut néanmoins se méfier des conclusions hâtives. Par exemple, l'effet de la demande en travail d'un atelier sur la fatigue et le stress d'un agriculteur, n'a de sens que si celle-ci est remise en perspective avec la demande globale en travail sur l'ensemble de l'exploitation. La dimension temporelle des effets est également différente de celle des données collectées relatives aux flux entrants de l'étape 1 car il existe bien souvent un décalage temporel entre les causes et les effets, certains effets sont immédiats, d'autres sont plus longs à observer. Par exemple, l'impact d'un système d'élevage sur l'état de santé d'une population ou le changement climatique se manifeste à long terme, alors que les nuisances (odeurs, bruits...) occasionnées dans le voisinage d'une exploitation d'élevage sont instantanées.

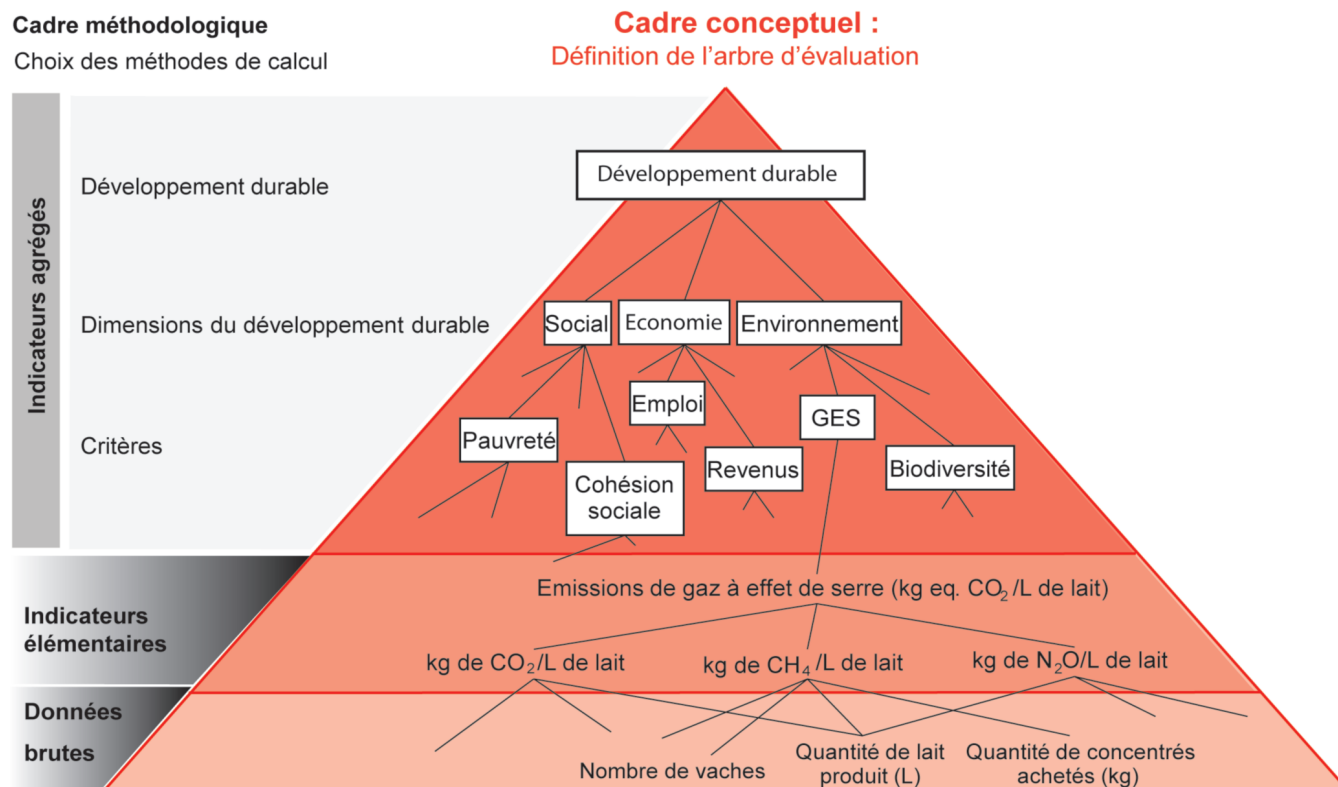
### 1.3 / Définition du cadre conceptuel et méthodologique

Comme évoqué précédemment, le développement durable est un concept polymorphe. Les évaluations des contributions de l'agriculture au développement durable résultent de nombreux choix : choix des critères, de leur organisation et de leur importance relative, choix des indicateurs pertinents pour rendre compte des critères et des valeurs de référence pour les interpréter.

Les critères pris en compte peuvent varier d'une évaluation à l'autre, le résultat final insiste donc plus ou moins sur une problématique donnée suivant les critères considérés par la méthode. La manière dont est traitée une même problématique peut également varier. Par exemple la problématique de l'azote en élevage englobe de nombreux niveaux : l'efficacité de rétention de l'azote par les animaux, les émissions associées à la production des intrants, la gestion et le stockage des effluents et leurs émissions associées, la lixiviation du nitrate au champ après épandage et les émissions gazeuses NH<sub>3</sub>, NO<sub>x</sub> associées, etc. Le choix des critères d'évaluation n'est donc pas anodin. De plus, pour un même critère, de nombreux indicateurs existent, ainsi qu'une multitude de valeurs de référence possibles pour les interpréter, qui peuvent être dépendantes du contexte local et modulées en fonction de la sensibilité du milieu (caractéristiques écologiques, biophysiques, économiques ou sociales) dans lequel l'évaluation est conduite.

#### a) *Le cadre conceptuel*

Il permet d'explicitier la vision du développement durable retenue (figure 3). Il fait le lien entre les dimensions du développement durable, leur déclinaison en critères, et l'élaboration et l'organisation des indicateurs. Son élaboration doit être claire et transparente de sorte que les indicateurs puissent être mis en relation avec les enjeux et objectifs (critères) et ainsi garantir que l'interprétation des résultats d'évaluation ne soit pas erronée. Deux types de démarches peuvent être utilisées pour définir le cadre conceptuel (Clément et Rey-Vaite 2008) : la démarche descendante (ou substantielle) ou la démarche ascendante (participative ou procédurale). La première repose sur des experts (recherche, institutions internationales) qui s'appuient sur des référentiels connus. L'intérêt de cette démarche est de profiter d'une normalisation antérieure, de cadres extérieurs robustes ayant fait leurs preuves, de faire référence à des normes internationales. Son risque réside dans la légitimité démocratique et l'appropriation par les acteurs en cas de non-adaptation aux réalités locales.

**Figure 3.** Cadres conceptuels et méthodologiques en évaluation multicritère appliqué à l'élevage laitier.

Cadre conceptuel : vision du développement durable déclinée en dimensions, critères et indicateurs.

Cadre méthodologique : organisation des calculs (comment un indicateur est calculé, comment l'agrégation est réalisée).

La seconde démarche repose sur les acteurs eux-mêmes, qui se constituent en collectif. Cette démarche est souvent initiée à partir d'une demande sectorielle, territoriale ou pour un but de certification. Elle présente l'avantage d'être basée sur les perceptions locales des problèmes et donc d'être plus facilement appropriable par les acteurs. Le risque associé est celui d'une faible validité en raison d'un accès limité de ces acteurs aux connaissances scientifiques sur les systèmes et les processus en jeu, ou au manque de moyens pour les utiliser. À l'inverse, une méthode descendante peut aussi porter le risque d'une méconnaissance de la complexité des problématiques locales. Se pose également la question de la comparabilité des résultats et de la reconnaissance par des parties prenantes non associées. Les démarches participatives ne sont pas simples à conduire et présentent un certain nombre d'écueils. Slocum *et al* (2006) ont présenté les modalités de mise en œuvre de ces démarches participatives. Des approches combinant préoccupations des acteurs et prise en compte des connaissances scientifiques, basées sur des échanges entre chercheurs et acteurs locaux sont aussi envisageables.

#### b) Le cadre méthodologique

Il définit l'ensemble des choix méthodologiques essentiels permettant de passer des données disponibles aux résultats obtenus (ex. choix des valeurs de réf-

rence, transformation et normalisation des variables, pondération, sélection des modes d'agrégation et des éventuels niveaux intermédiaires d'agrégation, post-analyse (analyse de sensibilité), présentation des résultats). Sa transparence garantit la reproductibilité de la méthode (figure 3). Une partie clé de ce cadre méthodologique : l'agrégation, est développée dans la partie 3.

## 2 / Choisir ou construire sa méthode

### 2.1 / Définir sa stratégie d'évaluation

Conduire une évaluation multicritère peut consister à avoir recours à une méthode existante ou bien à développer une nouvelle méthode. Chacune de ces stratégies d'évaluation présentent des avantages et des inconvénients. Par exemple l'utilisation d'une méthode existante peut permettre de réaliser des comparaisons entre différents systèmes d'élevage précédemment évalués ou d'échanger entre évaluateurs utilisant la même méthode. En revanche, il est justifié de construire sa propre méthode d'évaluation lorsque les méthodes existantes ne prennent pas en compte les besoins ou les contraintes des acteurs concernés. Le choix de construire une autre méthode peut aussi être motivé par le fait que les cadres conceptuels existants sont

trop restreints ou trop éloignés de celui des acteurs, que les indicateurs et les valeurs de références ne conviennent pas, ou encore que la méthode d'agrégation des indicateurs, si elle est nécessaire, ne correspond pas à ce qui est souhaité par les acteurs et/ou concepteurs. Construire sa propre méthode c'est aussi offrir aux groupes sociaux concernés par le système évalué l'opportunité de partager leur vision du développement durable (développement du cadre conceptuel) et de s'approprier la méthode d'évaluation. Par contre, le processus est chronophage et plus complexe, il nécessite de suivre une certaine démarche afin d'assurer une cohérence globale de l'évaluation.

Avant de s'engager dans une stratégie d'évaluation particulière et afin d'éviter de faire un choix inadapté, il est indispensable de s'interroger sur un certain nombre d'éléments qui vont conditionner la démarche à adopter.

### 2.2 / Utiliser une méthode existante

Il est important que le choix d'une méthode ne soit pas uniquement dicté par des conditions contingentes (accessibilité, simplicité...), mais surtout par sa pertinence vis-à-vis des objectifs de l'évaluation (Bockstaller *et al* 2013).

Lors du choix d'une méthode, il est nécessaire de vérifier la compatibilité des

finalités de l'évaluation avec celles annoncées par les méthodes existantes. Ces finalités des méthodes existantes sont nombreuses (cf. § 1.1), quelques questions peuvent guider le choix, faisant référence aux étapes A et B de la figure 2 :

*i) Quelle est l'audience ou le public visé par la méthode ?*

Chaque méthode a été construite pour un public particulier. Il peut s'agir d'agriculteurs pour du conseil individuel, de groupes d'acteurs d'une filière, de décideurs politiques, etc. Identifier précisément l'audience et le public visé par la méthode permet de s'assurer que les termes employés par la méthode, et la manière dont sont présentés les résultats sont adaptés.

*ii) Quelle est la facilité de mise en œuvre et de prise en main de la méthode par l'applicateur ?*

Certaines méthodes sont conçues pour être vulgarisées à un très large public, d'autres nécessitent un haut niveau d'expertise et peuvent impliquer le recours à des données difficiles d'accès.

*iii) Quels sont les enjeux qui ont guidé la création de la méthode ?*

Il est important de connaître le contexte spécifique dans lequel la méthode a été créée afin de s'assurer que sa réutilisation est adaptée et non biaisée par des objectifs particuliers des concepteurs.

*iv) La vision du développement durable sous-tendue par la méthode est-elle compatible avec les enjeux définis dans l'évaluation ?*

*v) Quel est le système évalué par la méthode ?*

La plupart des méthodes n'explicitent pas clairement les limites du système évalué (cf. § 1.2).

*vi) Quelles sont les bases scientifiques et méthodologiques ?*

Les cadres conceptuel et méthodologique de la méthode identifiée doivent être compatibles avec ceux de l'évaluation (cf. § 2.3).

## 2.3 / Construire une méthode

### a) La démarche

Après avoir précisé l'objet de l'évaluation (les finalités et le système évalué) et clarifié les contraintes (temps disponible, ressources financières et humaines, type d'informations nécessaires et disponibles, niveau de complexité des résultats,

format et représentation des résultats...) pour construire sa propre méthode d'évaluation, il faut définir (figure 4) :

*i) le cadre conceptuel qui définit la vision du développement durable sous-tendue par la méthode d'évaluation et le lien entre les dimensions du développement durable, les critères, et les indicateurs (cf. § 1.3) ;*

*ii) le cadre méthodologique qui définit l'ensemble des choix qui permettent de passer des données disponibles aux résultats obtenus (c.f. § 1.3 et partie 3) ;*

*iii) le mode de représentation des résultats, qui peut couvrir différentes formes.*

### b) Les points de vigilance à respecter

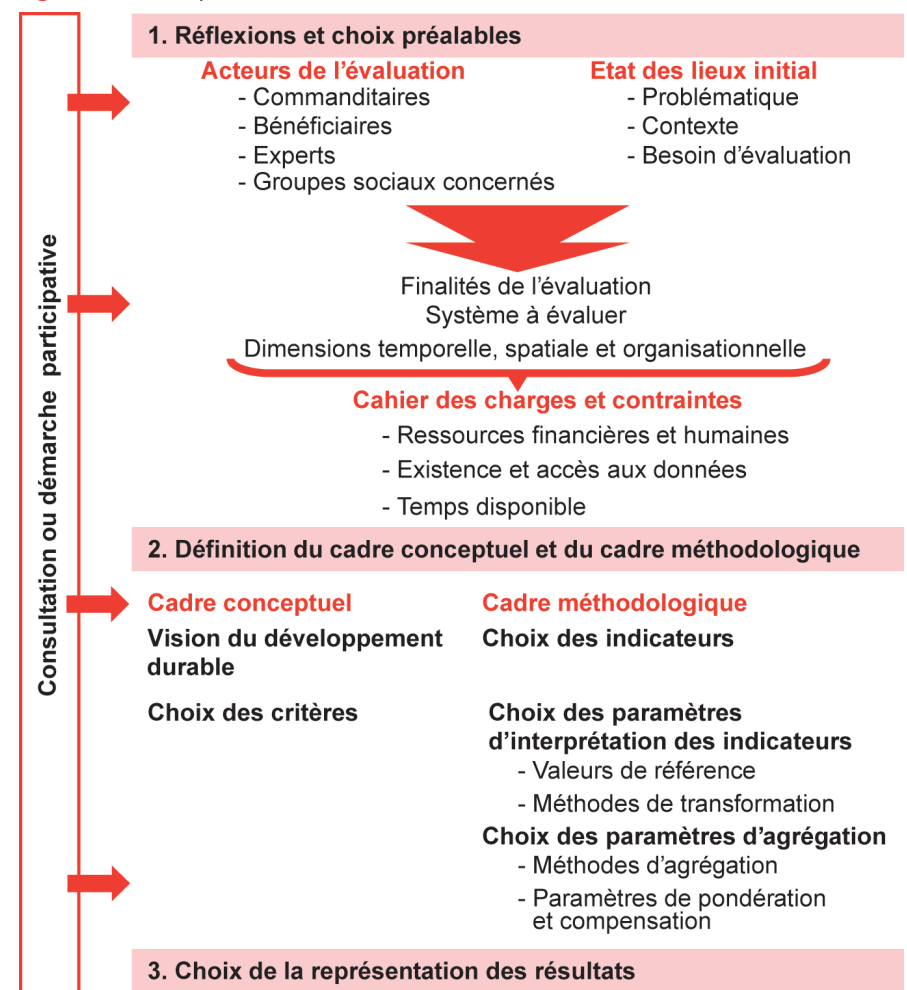
Cette démarche de conception d'une nouvelle méthode est délicate et nécessite de la rigueur. Plusieurs points de vigilance doivent être gardés à l'esprit.

*Être clair sur les contraintes de l'évaluation.* Il est nécessaire de considérer les contraintes suivantes : le temps et les ressources financières et humaines (compétences, effectif) disponibles pour construire et réaliser l'évaluation, ainsi que

les données (existence, disponibilité, qualité). Ces contraintes n'affectent pas les finalités mais elles façonnent l'évaluation en ayant une influence sur le périmètre d'évaluation et le choix des indicateurs. En effet, les ressources vont influencer l'étendue (dimension spatiale et nombre d'effets pris en compte) et le grain (niveau de détail de l'information) de l'évaluation. De même, le type d'effet à évaluer nécessite des indicateurs de natures différentes qui sont plus ou moins complexes et qui requièrent donc plus ou moins de ressources et de compétences.

*Être transparent.* Il faut que les choix réalisés lors de la conception de la méthode soient transparents et clairement explicités. La transparence concerne notamment *i) le périmètre de l'évaluation, ii) les modalités de passage des données brutes en indicateurs et, éventuellement, de leur transformation en scores, et iii) des choix réalisés en cas d'agrégation des données (cf. partie 3).* En effet, l'agrégation s'accompagne généralement de transformation des indicateurs, de pondération et de compensation entre critères, qui ont une forte influence sur les résultats et rendent cette étape subjective. Le résultat va dépendre des objectifs du groupe concep-

Figure 4. Les étapes nécessaires à l'élaboration d'une méthode d'évaluation.



teur et cette étape de l'agrégation doit donc être totalement transparente pour l'utilisateur de la méthode.

*Être complet.* Il est nécessaire de vérifier que la méthode évalue l'ensemble des objectifs/critères définis dans le cadre conceptuel, même s'il est parfois difficile de choisir un (des) indicateur(s) qui présente(nt) les qualités requises (cf. encadré 4). Si l'exhaustivité ne peut être atteinte, il est alors nécessaire de documenter les choix réalisés pour que l'utilisateur soit informé et qu'il prenne du recul sur les résultats qu'il a obtenus.

*Être sensible mais robuste.* Il est important d'avoir un équilibre entre ces deux qualités complémentaires. D'un côté, la méthode doit mesurer correctement les modifications attendues (méthode sensible). C'est-à-dire qu'une modification de la valeur des indicateurs dans leur gamme de variations au sein du (des) système(s) évalué(s) doit se traduire dans le résultat de l'évaluation, et doit donc changer la conclusion de l'évaluation. Cette qualité permettra de distinguer les systèmes au regard de leur performance de durabilité, d'évaluer les marges de progrès, etc. D'un autre côté, une légère modification des données brutes utilisées, correspondant par exemple à leur gamme d'incertitude, ne doit modifier que très peu la conclusion de l'évaluation (méthode robuste). Notons, qu'il est parfois préférable d'utiliser des échelles pour l'attribution des scores d'indicateurs qui sont moins sensibles mais plus robustes. Par exemple, il peut dans certains cas être plus pertinent d'utiliser une échelle discrète plutôt que continue, ou d'utiliser une échelle avec moins de niveaux ; mais le risque d'avoir des effets de seuils importants est alors à prendre en considération.

*Éviter la redondance.* Il faut éviter que certains critères ne soient représentés plusieurs fois dans le cadre conceptuel, et que le même indicateur ne soit utilisé plusieurs fois pour évaluer des critères différents. Il faut aussi parfois éliminer des critères ou indicateurs trop corrélés entre eux. En effet, ces redondances peuvent donner plus de poids à une composante particulière du développement durable et biaiser les résultats de l'évaluation. Par exemple, le taux de mortalité des animaux est généralement une donnée facile d'accès en élevage et est un indicateur pertinent pour évaluer certaines dimensions dans le pilier économique, environnemental et social. En effet, le taux de mortalité affecte directement la rentabilité des ateliers puisque le chiffre d'affaire est proportionnel aux produits vendus. Mais, il est également lié à l'ensemble des impacts environnementaux car la mortalité représente des intrants qui sont consommés « inutilement » sans production finale. Dans le pilier social, le

#### Encadré 4. Qu'est-ce qu'un indicateur ?

Un indicateur est une variable mesurable qui fournit des renseignements sur une autre variable plus difficile d'accès (Gras *et al* 1989). Un indicateur peut être de nature très diverse : quantitatif (cardinal) ou qualitatif (nominal ou ordinal) ; descriptif ou prédictif ; élémentaire (constitué de données brutes) ou élaboré (calculé à partir de plusieurs données brutes) ; mesuré/estimé (calcul simple ou à partir d'un modèle) ou observé/déclaré (indicateur subjectif).

Un bon indicateur doit remplir les qualités suivantes selon les finalités de l'évaluation : transparent et clairement défini (méthode, source des données, hypothèses sous-jacentes, périmètre) ; validé et reconnu scientifiquement (bases théoriques solides) et statistiquement (précis, fiable, robuste) ; reproductible dans des contextes différents (comparabilité dans le temps et l'espace) ; sensible au changement lorsque le système change évolue ou lorsqu'il y a une rupture d'équilibre ; couvrant le problème qu'il mesure et mesurant réellement ce qu'il est supposé détecter ; évite les mauvaises prises de décisions ; facile à renseigner et à interpréter (accessibilité des données) ; simple à calculer (temps et coût de mise en œuvre) ; adapté aux utilisateurs visés (niveau de qualification, réponse aux besoins, résultat compréhensible, facilement communicable).

taux de mortalité est généralement utilisé comme indicateur du bien-être animal. Malgré ses qualités, pertinence et facilité d'accès, l'utilisation de ce même indicateur dans chacune des trois dimensions reviendrait à affaiblir l'évaluation en limitant plutôt qu'en élargissant l'analyse du système. Le choix d'utiliser ou non plusieurs fois le même indicateur doit alors être le résultat d'un compromis entre redondance et pertinence. Il peut arriver qu'un même indicateur puisse être utilisé plusieurs fois dans une évaluation mais avec une interprétation différente en fonction des critères auxquels ils se rapportent.

*Valider la méthode d'évaluation.* La validation de la méthode concerne différents points et peut être réalisée de plusieurs manières. Elle peut engager un groupe d'experts extérieurs au groupe des concepteurs ou mobiliser un groupe d'acteurs concernés par la méthode qui vérifie si la méthode respecte les points précédents (transparence, complétude, sensibilité, robustesse, redondance). Il est aussi possible de réaliser une ou plusieurs évaluations test sur des cas connus pour vérifier *i)* si la méthode respecte les contraintes identifiées au départ (temps nécessaire, budget, qualification humaine, complexité, format et représentation des résultats), *ii)* si les indicateurs ont les qualités requises (mesurables, sensibles, pertinents...) en tenant compte de la situation actuelle et des changements prévisibles et *iii)* si (le cas échéant) la transformation des données en scores couvre l'amplitude des réponses et conserve une variabilité suffisante et pertinente.

### 3 / Les enjeux de l'agrégation

En fonction de la finalité de l'évaluation, l'agrégation formalisée des indicateurs puis des critères, qu'elle soit

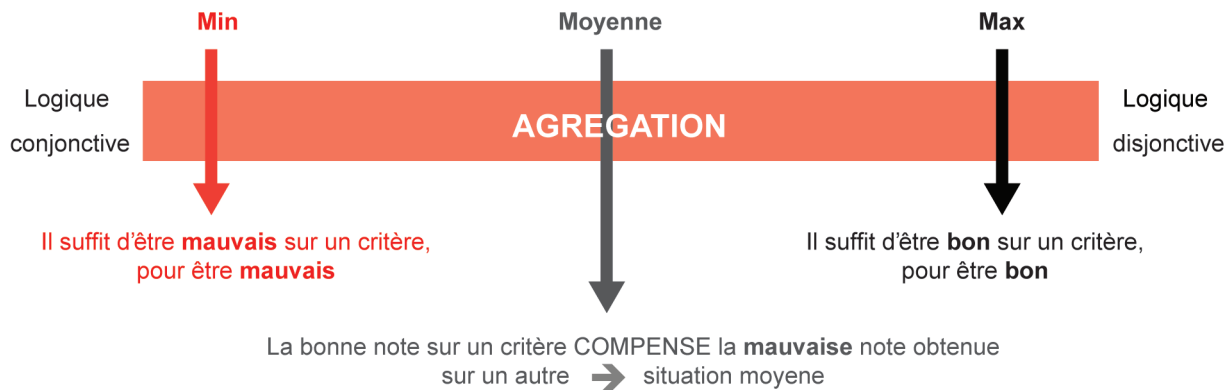
complète (jusqu'à une note globale de durabilité) ou non, peut s'avérer nécessaire. Par exemple, dans le cadre d'une procédure de certification, l'évaluation doit finir par conclure sur la labellisation ou non de l'exploitation ou du produit. Il s'agit d'une agrégation complète conduisant à une réponse binaire : labellisé vs non labellisé. Dans le cas où l'agrégation formalisée n'est pas nécessaire et où une simple évaluation descriptive est requise, l'utilisateur final va généralement utiliser les résultats désagrégés pour émettre un jugement global. Par exemple, dans le cadre d'une évaluation pour aider à concevoir des systèmes innovants, une conclusion sera formulée sur le système innovant le plus performant. Cette conclusion est en fait le fruit d'une agrégation complète, mais non formalisée puisque réalisée par l'utilisateur sur la base de l'évaluation descriptive. Les étapes de raisonnement suivies pour réaliser cette agrégation ne sont généralement pas explicitées, rendant l'évaluation peu transparente. Dans tous les cas, que l'agrégation soit formalisée (*via* un calcul mathématique) ou non (à dire d'experts sur la base de l'évaluation descriptive), cette dernière repose sur deux raisonnements différents et complémentaires : d'une part la gestion des pondérations, et d'autre part celle des compensations entre critères.

Les pondérations interviennent quand un critère est jugé plus important qu'un autre. Par exemple, si on évalue la durabilité environnementale d'un système d'élevage, la préservation de la biodiversité est-elle plus ou moins importante que la limitation du réchauffement climatique ? Il s'agit du raisonnement que nous avons tous instinctivement en tête, même si répondre à la question n'est pas forcément évidente. Pour ce faire, plusieurs cas de figure existent. Tout d'abord, l'utilisation de connaissances scientifiques

**Tableau 1.** Méthode OVALI – exemple des pondérations entre les trois piliers de la durabilité et les critères qui les composent établies par les acteurs en nombre de points et retranscrites ici en % pour plus de lisibilité.

| Pilier          |        |           | Critères  |        |           |
|-----------------|--------|-----------|---|--------|-----------|
| Nom             | Points | Poids (%) | Nom   | Points | Poids (%) |
| Economique      | 180    | 33        | Créer de la valeur sur le territoire                | 73     | 41        |
|                 |        |           | Connecter les filières au marché                    | 67     | 37        |
|                 |        |           | Participer à l'autosuffisance alimentaire française | 40     | 22        |
| Social          | 180    | 33        | Répondre aux attentes des citoyens                  | 84     | 47        |
|                 |        |           | Favoriser l'acceptabilité sociale de la filière     | 62     | 34        |
|                 |        |           | Renforcer le lien avec le territoire                | 34     | 19        |
| Environnemental | 180    | 33        | Optimiser la gestion des ressources                 | 69     | 38        |
|                 |        |           | Maîtriser les impacts environnementaux              | 68     | 38        |
|                 |        |           | Préserver les milieux naturels sur les sites        | 43     | 24        |

**Figure 5.** Les différentes logiques de compensation, de la plus sévère à la plus généreuse.



permet dans certains cas de mettre en évidence des équivalences entre les différents éléments à agréger. C'est la démarche qui est par exemple utilisée pour pondérer les émissions de CO<sub>2</sub>, de CH<sub>4</sub> et de N<sub>2</sub>O pour leur contribution au réchauffement climatique, avec des poids respectifs de 1, 25 et 298 qui sont utilisés pour calculer l'indicateur sur la contribution au changement climatique (en équivalent CO<sub>2</sub>). Cette approche ne peut cependant s'appliquer qu'à l'agrégation d'indicateurs (et non à l'agrégation de critères), et même plus précisément d'indicateurs pour lesquels il est scientifiquement possible d'établir des équivalences. Une autre démarche consiste à se baser sur des analyses statistiques, notamment d'analyse des données (ex. Analyse en Composantes Principales, ACP) pour déterminer les poids de différents facteurs dans le calcul d'un indicateur. Dans le cas d'une ACP, la part de variabilité observée sur le premier axe peut servir à pondérer les différents indicateurs. Les pondérations sont alors uniquement dépendantes de la variabilité des indicateurs observée sur le jeu de données. Cette démarche peut être discutable, et les pondérations ainsi obtenues sont valables uniquement sur le jeu de données utilisé. Enfin, la dernière démarche pouvant être utilisée consiste à se baser sur

l'opinion d'experts, de décideurs ou de groupes sociaux concernés par le système évalué. Dans cette approche, les poids peuvent être attribués en référence aux réglementations en vigueur, aux objectifs des agriculteurs ou encore aux enjeux du territoire. Dans la méthode OVALI d'évaluation de la filière avicole, une pondération permet de refléter l'importance accordée par le groupe participatif aux différents objectifs, critères et indicateurs. Ainsi, le choix a été fait d'attribuer un même nombre de points à chacun des trois piliers de la durabilité (180 points), de façon à pouvoir apprécier la stabilité du « trépied » pour le système étudié. Ces 180 points sont ensuite répartis entre les objectifs puis les critères et les indicateurs (Dusart *et al* 2015, voir tableau 1 pour les pondérations des piliers et des critères).

Les réflexions sur les compensations entre critères sont généralement moins fréquentes et pourtant elles sont tout aussi importantes dans la procédure d'évaluation. La gestion des compensations consiste à savoir jusqu'à quel point nous souhaitons compenser la mauvaise performance obtenue sur un critère par une bonne performance sur un autre. Lors de l'évaluation, les compensations peuvent être « thématiques » (par exemple, pour la biodiversité, le mauvais niveau de bio-

diversité domestique peut-il être compensé par le bon niveau de biodiversité sauvage sur l'exploitation ?), « spatiales » (par exemple à l'échelle d'un territoire : les mauvaises performances de quelques exploitations sont-elles compensées par les bonnes performances des autres ?), ou encore « temporelles » (par exemple les mauvaises performances économiques d'une exploitation une année peuvent-elles être compensées par les performances des années précédentes et suivantes ?). Différentes logiques de compensations existent (figure 5), de la plus sévère, la logique conjonctive, dans laquelle il faut être bon partout pour être bon au global, à la plus souple, la logique disjonctive, dans laquelle il suffit d'être bon quelque part pour être bon au global. En fonction de ce qu'on cherche à évaluer, il faut donc identifier où l'on souhaite se positionner sur ce continuum. Pour ce faire, la démarche consiste à se baser sur l'opinion d'experts, de décideurs ou de groupes sociaux concernés. Dans le cas de l'évaluation de la durabilité des systèmes d'élevage, la pondération des piliers économique, social et environnemental va dépendre du point de vue des experts ou acteurs consultés, entre « durabilité forte » (où le capital naturel n'est pas substituable par du capital manufacturé et où les compensa-

tions sont interdites) et « durabilité faible » (où les deux types de capital sont substituables) (Ayres *et al* 1998).

Les raisonnements sur les pondérations et les compensations sont parfaitement complémentaires. Le tableau 2 présente l'exemple du modèle d'évaluation du bien-être des poulets de chair, développé dans le cadre du projet Welfare Quality® (Welfare Quality® consortium 2009), et plus précisément l'agrégation des critères « Absence de faim » et « Absence de soif » pour évaluer le principe « Bonne alimentation » (un des 4 principes de l'évaluation du bien-être). Dans ce projet, des experts ont été consultés pour paramétrer le modèle d'évaluation, c'est-à-dire établir les pondérations entre ces deux critères et définir le niveau de compensation à autoriser. Le tableau 2 présente le cas de trois exploitations, les notes qu'elles ont obtenues sur les deux critères ainsi que la note souhaitée par un des experts. On peut constater que cet expert a joué à la fois sur les pondérations et sur les compensations. En effet, en attribuant une note plus élevée à l'exploitation A qu'à l'exploitation C il considère que le critère « Absence de soif » est plus important que le critère « Absence de faim ». Par ailleurs, en attribuant une note moins élevée à l'exploitation A qu'à l'exploitation B il considère que la mauvaise note obtenue sur le critère « Absence de faim » par l'exploit-

tation A ne peut pas être compensée par la bonne note obtenue sur le critère « Absence de soif », pourtant le critère le plus important à ses yeux.

Le tableau 2 présente également à titre d'exemple les résultats qui ont pu être obtenus par optimisation en utilisant cinq méthodes de calcul différentes : la moyenne pondérée, qui ne joue que sur les pondérations, le minimum, le maximum et la moyenne pondérée ordonnée, qui ne jouent que sur les compensations, et l'intégrale de Choquet, qui joue à la fois sur les pondérations et les compensations. Il est facile de constater que dans cet exemple le choix de la méthode de calcul influence énormément les résultats, ce qui illustre la nécessité de choisir la méthode de calcul en fonction des raisonnements que l'on souhaite suivre lors de l'agrégation.

#### 4 / Quitter le paradigme de la science qui sait tout

Entreprendre une évaluation multicritère est une démarche subjective puisque le résultat dépend en partie des choix réalisés sur les objectifs et les priorités des acteurs de l'évaluation (Girardin *et al* 1999). Ces objectifs et priorités se traduisent par des choix tout au long de l'évaluation, par exemple le choix du périmètre d'évaluation, celui des critères et des

indicateurs et de leur pondération dans le cas d'une agrégation. L'évaluation multicritère utilise des indicateurs qui synthétisent ou simplifient des variables pour rendre compte d'un phénomène qui ne peut être décrit directement en raison de sa complexité ou pour des raisons de faisabilité. Elle interprète donc une réalité complexe en la simplifiant, ce qui peut s'apparenter également à de la subjectivité. Pour autant, cela ne signifie pas que les résultats d'une évaluation multicritère n'ont aucune valeur scientifique, et cela même s'il est souvent impossible de les vérifier par comparaison à une réalité observée sur le terrain (il n'existe par exemple aucune « mesure » de la durabilité à laquelle comparer les sorties d'une méthode d'évaluation de la durabilité pour valider le modèle). La portée scientifique de l'évaluation est assurée par la transparence des choix réalisés, la sensibilité et la robustesse de la méthodologie, la pertinence des critères évalués vis-à-vis des enjeux traités, et la cohérence du cadre conceptuel sur lequel l'évaluateur développe sa méthodologie. Bien que subjective, l'évaluation doit éviter toute appréciation partielle et arbitraire, mais plutôt trouver la solution la plus adéquate à un problème à plusieurs dimensions dans un contexte particulier.

Les processus biophysiques comme le taux de lixiviation du nitrate ou la quantité d'émissions de gaz à effet de serre

**Tableau 2.** Exemple d'agrégation de deux critères (« absence de faim », basé sur l'indicateur « % d'animaux émaciés », et « absence de soif », basé sur l'indicateur « nombre d'animaux par abreuvoir ») utilisés pour évaluer le principe « bonne alimentation » dans l'évaluation du bien-être des poulets de chair de Welfare Quality® (d'après Lairez *et al* 2015).

Les valeurs des critères et du principe sont exprimées sur une échelle entre 0 (la pire situation envisageable vis-à-vis du bien-être) et 100 (la situation idéale). Pour aller plus loin consulter : Yager (1988) sur les moyennes pondérées ordonnées et Mayag *et al* (2011) sur l'intégrale de Choquet.

| Exploitation   | Critère « Absence de faim » | Critère « Absence de soif » | Principe « Bonne alimentation » |  |                        |                  |                           |                      |                  |                |
|--|-----------------------------|-----------------------------|---------------------------------|--|------------------------|------------------|---------------------------|----------------------|------------------|----------------|
|  |                             |                             | Note souhaitée par l'expert     | Note calculée par différentes méthodes |                        |                  |                           |                      |                  |                |
|  |                             |                             |                                 | Minimum des 2 critères                 | Maximum des 2 critères | Moyenne pondérée | Moyenne pondérée ordonnée | Intégrale de Choquet |                  |                |
| <b>A</b>   | 25                          | 75                          | <b>40</b>                       | 25                                     | 75                     | 55               | 35                        | 40                   |                  |                |
| <b>B</b>   | 50                          | 50                          | <b>50</b>                       | 50                                     | 50                     | 40               | 50                        | 50                   |                  |                |
| <b>C</b>   | 75                          | 25                          | <b>30</b>                       | 25                                     | 75                     | 45               | 35                        | 30                   |                  |                |
| Somme des carrés des écarts entre la note souhaitée par l'expert et la note calculée par les différentes méthodes  |                             |                             |                                 | 250                                    | 3250                   | 450              | 50                        | 0                    |                  |                |
| Valeur des paramètres nécessaires aux calculs avec :   |                             |                             |                                 | -                                      | -                      | $w_1 = 0,4$      | $\omega_1 = 0,8$          | $w_1 = 0,4$          |                  |                |
|  |                             |                             |                                 |  |                        |                  |                           | $w_2 = 0,6$          | $\omega_2 = 0,2$ | $w_2 = 0,6$    |
|  |                             |                             |                                 |  |                        |                  |                           | -                    | -                | $I_{12} = 0,6$ |
| <ul style="list-style-type: none"> <li>• <math>w_1</math> : le poids du critère « Absence de faim »</li> <li>• <math>w_2</math> : le poids du critère « Absence de soif »</li> <li>• <math>\omega_1</math> : le poids du plus mauvais critère</li> <li>• <math>\omega_2</math> : le poids du meilleur indicateur</li> <li>• <math>I_{12}</math> : valeur de l'indice d'interaction entre les deux critères (plus cette valeur est importante plus les compensations entre les deux critères sont limitées, cf. logique conjonctive)</li> </ul> |                             |                             |                                 |  |                        |                  |                           |                      |                  |                |

d'un troupeau sont souvent intégrés dans les évaluations multicritères des systèmes d'élevage. Les indicateurs utilisés pour quantifier ces processus biophysiques reposent sur des bases scientifiques solides. En revanche, lorsqu'il s'agit d'agréger un ensemble d'informations pour aboutir à une évaluation des contributions au développement durable, la question de la subjectivité se pose. En effet, la durabilité est avant tout un construit social, avec de nombreuses visions possibles. Comment choisir ce qui a le plus d'importance dans les enjeux du développement durable ? Dans l'évaluation, allons-nous privilégier la croissance économique à la protection de l'environnement, ou encore, la compétitivité des filières d'élevage au bien-être des travailleurs ? Faut-il attribuer le même poids à la préservation de la qualité de l'eau et à l'utilisation d'énergies non-renouvelables ? La réponse à ces questions n'est pas toujours facile, c'est pourquoi l'évaluation multicritère de la durabilité peut être utilisée comme un objet intermédiaire permettant la concertation entre acteurs. Le dixième principe de la déclaration issue du sommet de Rio de 1992 affirme même que « la meilleure façon de traiter les questions d'environnement est d'assurer la participation de tous les citoyens concernés ».

De nombreuses méthodes participatives ont été développées pour concerter des acteurs autour de l'évaluation de la durabilité. À titre d'exemple, la méthode OVALI (outil multicritère d'évaluation de la durabilité pour concevoir des systèmes avicoles innovants) a réuni l'ensemble des acteurs d'une filière de poulet de chair afin d'identifier des points forts et des points faibles, et de proposer des voies de progrès pour améliorer la durabilité de la filière (Protino et Bouvarel, 2013). La méthode est utilisée comme un outil d'accompagnement et d'aide à la décision pour l'ensemble des maillons d'une filière intégrée. Le groupe participatif était constitué d'acteurs concernés par la durabilité de la filière : les éleveurs, les abatteurs, les transformateurs, les consommateurs, etc. Le groupe a été consulté pour la définition des objectifs du développement durable, des critères, des indicateurs et de leur transformation, et de la pondération. Un autre exemple est la méthode EVAD (ÉVALuation de la Durabilité des systèmes de production

aquacoles). EVAD est une méthode générique d'analyse de la durabilité de l'aquaculture, pouvant s'adapter aux différents types de systèmes aquacoles. Elle propose une grille d'indicateurs de la durabilité et apporte un diagnostic analysé avec les acteurs, à l'échelle du territoire (Rey-Valette *et al* 2008). Un des postulats de cette méthode est que la démarche de co-construction d'indicateurs facilite la concertation entre acteurs et favorise l'apprentissage organisationnel. Ces approches participatives impliquent les acteurs concernés par la définition des enjeux de l'évaluation, et dans le choix des développements méthodologiques. De ce fait, la méthode résulte d'un consensus basé sur leurs priorités à l'égard du développement durable. Les scientifiques se positionnent alors comme des accompagnateurs/facilitateurs, ils donnent les informations utiles à la prise de décision par les participants.

## Conclusion

Les systèmes d'élevage sont en évolution permanente. Évaluer les pratiques et leur évolution est un enjeu important de développement et de gouvernance. Les différents groupes d'acteurs sont en attente de réponses à de nombreuses échelles (ateliers, exploitations, territoires, filières) pour guider et accompagner les différents secteurs des productions animales. L'évaluation n'est plus réservée à des observateurs extérieurs mais implique de plus en plus les producteurs et les acteurs de terrain. L'évaluation nécessite de la rigueur dans la conception et la conduite pour conserver sa pertinence et sa crédibilité, d'autant plus s'il y a une multiplicité d'acteurs. L'évaluation multicritère doit donc bénéficier d'un cadre et d'un ensemble de règles transparentes et partagées. Pour construire ce cadre, il faut organiser la nécessaire collaboration entre recherche, instituts techniques, acteurs du développement et acteurs de terrain. Par ailleurs dans la conception des outils d'évaluation multicritère des systèmes de production animale, les écueils sont nombreux et ils s'étendent depuis la création d'outils orientés pour défendre un point de vue partisan, jusqu'à la perte de vue des enjeux en se noyant dans la technique. L'utilisateur final de l'évaluation ne doit pas être oublié au risque de concevoir de jolis outils qui ne

servent pas. Pour ce faire, concepteurs et utilisateurs ont besoin d'être accompagnés dans leurs démarches d'évaluation. Des initiatives existent pour associer ces compétences et faire converger les démarches d'évaluation multicritères, par exemple le Réseau Mixte Technologique Erytage<sup>1</sup>, ou au sein des Groupements d'intérêt scientifiques comme le GIS Élevage demain<sup>2</sup>. Des outils mutualisés comme la plateforme MEANS (<https://www.inra.fr/means>) pourraient être aussi mobilisés.

Il est nécessaire de poursuivre les études sur l'évaluation multicritère appliquée aux systèmes agricoles, et de les faire reconnaître comme des démarches scientifiques à part entière en s'appuyant sur les théories développées en mathématique et en informatique. Il est nécessaire de sortir de l'empirisme de la construction d'une liste d'indicateurs dans l'urgence et sans recul méthodologique. Le rapprochement avec la recherche opérationnelle notamment sur les méthodes d'agrégation et sur le traitement de l'information sont une des pistes à poursuivre.

De nombreuses questions scientifiques existent et se mêlent ; il convient donc de différencier celles qui portent sur l'évaluation, sur les systèmes, ou encore sur la notion même de développement durable. En particulier les indicateurs, au cœur de l'évaluation multicritère sont un champ d'investigation large. Un des enjeux est peut-être de combiner différentes qualités d'indicateurs, entre ceux qui sont robustes, reconnus et généralisables, et ceux qui sont spécifiques à des contextes particuliers. Il faut aussi être capable de mobiliser les outils mathématiques adaptés permettant une interprétation des résultats (définition de valeurs de référence, choix de la méthode d'agrégation...), et donc développer des compétences adaptées.

## Remerciements

Les auteurs remercient le GIS Élevage Demain et le GIS Grandes cultures HP2E, pour leur soutien dans le déroulement des travaux préalables à cet article et la rédaction du guide méthodologique sur l'évaluation multicritère (Lairez *et al* 2015).

<sup>1</sup> <http://www.plage-evaluation.fr/webplage/>

<sup>2</sup> <https://www.gis-elevages-demain.org/>

## Références

- Ayres R.U., van den Bergh J.C.J.M., Gowdy J.M., 1998. Viewpoint: Weak Versus Strong Sustainability, Tinbergen Institute Discussion Papers. Tinbergen Institute: Amsterdam, NL.
- Belhouchette H., Louhichi K., Therond O., Mouratiadou I., Wery J., Ittersum M.V., Flichman G., 2011. Assessing the impact of the Nitrate Directive on farming systems using a bio-economic modelling chain. *Agric. Syst.*, 104, 135-145.
- Bockstaller C., Guichard L., Makowski D., Aveline A., Girardin P., Plantureux S., 2008. Agri-environmental indicators to assess cropping and farming systems. A review. *Agron. Sustain. Dev.*, 28, 139-149.
- Bockstaller C., Guichard L., Keichinger O., Girardin P., Galan M.B., Gaillard G., 2009. Comparison of methods to assess the sustainability of agricultural systems. A review. *Agron. Sustain. Dev.*, 29, 223-235.
- Bockstaller C., Cariolle M., Galan M.B., Guichard L., Leclercq C., Morin A., Surleau-Chambenoit C., 2013. Évaluation agri-environnementale et choix des indicateurs : acquis, enjeux et pistes. *Innov. Agron.*, 31, 1-14.
- Buclet N., 2011. Écologie industrielle et territoriale, stratégies locales pour un développement durable. Presses Universitaires du Septentrion, 309p.
- Clément O., Rey-Valette H., 2008. Analyse des démarches de construction des indicateurs de développement durable, école chercheur. L'évaluation du développement durable, Sainte Foy-lès-Lyon, France.
- Dong Z., Pan Z., An P., Wang L., Zhang J., He D., Han H., Pan X., 2015. A novel method for quantitatively evaluating agricultural vulnerability to climate change. *Ecol. Indicators*, 48, 49-54.
- Dumont B., Fortun-Lamothe L., Jouven M., Thomas M., Tichit M., 2012. Prospects from agroecology and industrial ecology for animal production in the 21<sup>st</sup> century. *Animal*, 1-16.
- Dusart L., Protino J., Meda B., Dezat E., Chevalier D., Magdelaine P., Chenut R., Ponchant P., Lescoat P., Berri C., Bouvarel I., 2015. OVALI, une méthode pour améliorer la durabilité des filières avicoles. *TeMA*, 4-12.
- Erkman S., 2004. Vers une écologie industrielle, comment mettre en pratique le développement durable dans une société hyper-industrielle. Charles Léopold Mayer, 251p.
- FAO, 2015. L'Après 2015 et les ODD. Nourrir les êtres humains, prendre soin de la planète. [http://www.fao.org/fileadmin/user\\_upload/mdg/doc/Sustag\\_fr-1.pdf](http://www.fao.org/fileadmin/user_upload/mdg/doc/Sustag_fr-1.pdf)
- García-Martínez A., Olaizola A., Bernués A., 2008. Trajectories of evolution and drivers of change in European mountain cattle farming systems. *Animal*, 3, 152-165.
- Giller K. E., Tittonell P., Rufino M. C., van Wijk M.T., Zingore S., Mapfumo P., Adjei-Nsiah S., Herrero M., Chikowo R., Corbeels M., Rowe E.C., Baijukya F., Mwijage A., Smith J., Yeboah E., van der Burg W.J., Sanogo O. M., Misiko M., de Ridder N., Karanja S., Kaizzi C., K'ungu J., Mwale M., Nwaga D., Pacini C., Vanlauwe B., 2011. Communicating complexity: Integrated assessment of trade-offs concerning soil fertility management within African farming systems to support innovation and development. *Agric. Syst.*, 104, 191-203.
- Girardin P., Bockstaller C., Werf H.V.d., 1999. Indicators: tools to evaluate the environmental impacts of farming systems. *J. Sustainable Agric.*, 13, 5-21.
- Gras R., Benoit M., Deffontaines J.P., Duru M., Lafarge M., Langlet A., Osty P.L., 1989. Le fait technique en agronomie. *Activité agricole, concepts et méthodes d'étude*, Institut National de la Recherche Agronomique, L'Hamarttan, Paris, France, 184p.
- Hansen J.W., 1996. Is agricultural sustainability a useful concept? *Agric. Syst.*, 50, 117-143.
- Koch P., Salou T., 2014. AGRIBALYSE : Rapport méthodologique - version 1.1, In: ADEME (Ed.), Angers, France, p386.
- Lairez J., Feschet P., Aubin J., Bockstaller C., Bouvarel I., Coord., 2015. Agriculture et développement durable, Guide pour l'évaluation multicritère. Educagri Éditions/Éditions Quae, Dijon/versailles, 226p.
- Landais E., Balent G., 1993. Introduction à l'étude des systèmes d'élevage extensif. In: E. Landais, G. Balent, dir., *Pratiques d'élevage extensif. Identifier, modéliser, évaluer. Études et Recherches sur les Systèmes Agraires et le Développement*, INRA Versailles, France, 27, 13-36.
- Laurent C., Hulin S., Agabriel C., Chassaing C., Botreau R., Monteils V., 2017. Co-construction of an assessment method of the environmental sustainability for cattle farms involved in a Protected Designation of Origin (PDO) cheese value chain, Cantal PDO. *Ecol. Indicat.*, 76, 357-365.
- Martin G., 2015. A conceptual framework to support adaptation of farming systems – Development and application with Forage Rummy. *Agric. Syst.*, 132, 52-61.
- Mayag B., Grabisch M., Labreuche C., 2011. A representation of preferences by the Choquet integral with respect to a 2-additive capacity. *Theory Decis.*, 71, 297-324.
- Millennium Ecosystem Assessment, 2005. *Ecosystems and human well-being: Synthesis*, In: Island Press (Ed). World Resources Institute, Washington, DC, 155p.
- Moraine M., Duru M., Nicholas P., Leterme P., Therond O., 2014. Farming system design for innovative crop-livestock integration in Europe. *Animal*, 8, 1204-1217.
- Protino J., Bouvarel I., 2013. Ovali : outil d'évaluation multicritère pour concevoir des systèmes de production avicoles innovants, INRA-ITAVI, Nouzilly, France, 235p.
- Reidsma P., Wolf J., Kanellopoulos, A., Schaap B.F., Mandryk M., Verhagen J., Van Ittersum M.K., 2015. Climate change impact and adaptation research requires integrated assessment and farming systems analysis: A case study in the Netherlands. *Environ. Res. Letters* 10, 4.
- Rey-Valette H., Clément O., Aubin J., Mathé S., Chia E., Legendre M., Caruso D., Mikolasek O., Blanchelon J.P., Slembrouck J., Baruthio A., René F., Levang P., Morissens P., Lazard J., 2008. Guide de co-construction d'indicateurs de développement durable en aquaculture, CIRAD-Ifrémer-INRA-IRD-UM1, Montpellier, France.
- Ryschawy J., Choisis N., Choisis J.P., Gibon A., 2012. Paths to last in mixed crop-livestock farming: lessons from an assessment of farm trajectories of change. *Animal*, 7, 673-681.
- Ryschawy J., Tichit M., Bertrand S., Allaire G., Plantureux S., Aznar O., Perrot C., Guinot C., Josien E., Lasseur J., Aubert C., Tchakerian E., Disenhaus C., 2015. Comment évaluer les services rendus par l'élevage ? Une première approche méthodologique sur le cas de la France. *INRA Prod. Anim.*, 28, 23-38.
- Schader C., Grenz J., Meier M.S., Stolze M., 2014. Scope and precision of sustainability assessment approaches to food systems. *Ecol. Soc.*, 19, 3.
- Slocum N., Elliott J., Heesterbeek S., Lukensmeyer C.J., 2006. Méthodes participatives, un guide pour l'utilisateur. Fondation Roi Baudouin, 202p.
- van der Werf H.M.G., Petit J., 2002. Evaluation of the environmental impact of agriculture at the farm level: a comparison and analysis of 12 indicator-based methods. *Agric. Ecosyst. Environ.*, 93, 131-145.
- Viaggi D., Raggi M., Gomez Y., Paloma S., 2013. Modelling and interpreting the impact of policy and price scenarios on farm-household sustainability: Farming systems vs. result-driven clustering. *Environ. Model. Software* 43, 96-108.
- Vilain L., Boisset K., Girardin P., Guillaumin A., Mouchet C., Viaux P., Zahm F., 2008. La méthode IDEA : indicateur de durabilité des exploitations agricoles : guide d'utilisation. 3<sup>ème</sup> édition, Educagri, Dijon, France, 184p.
- Welfare Quality® consortium, 2009. Welfare Quality® Assessment Protocol for Poultry, Lelystad, The Netherlands, 110 p. ISBN 978-90-78240-06-8
- Wolf J., Kanellopoulos A., Kros J., Webber H., Zhao G., Britz W., Reinds G.J., Ewert F., de Vries W., 2015. Combined analysis of climate, technological and price changes on future arable farming systems in Europe. *Agric. Syst.*, 140, 56-73.
- Yager R.R., 1988. On ordered weighted averaging aggregation operators in multicriteria decision-making. *IEEE Transactions on Systems, Man and Cybernetics*, 18, 183-190.
- Zahm F., Alonso Ugaglia A., Boureau H., Del'homme B., Barbier J.M., Gasselin P., Gafsi M., Guichard L., Loyce C., Manneville V., Menet A., Redlingshofer B., 2015. Agriculture et exploitation agricole durables : état de l'art et proposition de définitions revisitées à l'aune des valeurs, des propriétés et des frontières de la durabilité en agriculture. *Innov. Agron.*, 46, 105-125.

## Résumé

---

L'apparition de la notion de développement durable a conduit à un passage des évaluations orientées selon un seul critère vers des évaluations plus intégrées couvrant différentes dimensions de la performance des élevages : l'économie, le social, le sanitaire, le bien être, l'environnement. De nombreuses questions méthodologiques sont posées par la confrontation de ces différentes dimensions avec leurs échelles de mesure propres. Comment accompagner l'évolution des systèmes d'élevage avec une évaluation multicritère ? Comment bien définir les finalités de l'évaluation ? Comment gérer la multiplicité des informations ? Comment interpréter un ensemble d'indicateurs conjointement lorsqu'ils prennent des valeurs opposées ? Quels outils mobiliser ? Faut-il utiliser ou créer ses propres outils d'évaluation ? L'élevage avec ses enjeux multiples qui s'étagent entre les niveaux des ateliers, des filières et des territoires est particulièrement concerné par ces questions. La participation des acteurs dans la démarche d'évaluation est aussi une question à trancher, sont-ils co-concepteurs de la méthode, validateurs, enquêtés ou seulement destinataires des résultats ? Ces différents éléments conduiront à définir les meilleures méthodes de traitement mathématique des données. Ce traitement des données est intimement lié au mode de représentation des résultats et de leur utilisation. Cet article aborde les différents usages des évaluations multicritères appliquées aux systèmes d'élevage et les points méthodologiques clefs à respecter.

## Abstract

---

*The multicriteria assessment of livestock systems to support their evolution: issues, challenges and interrogations*

The emergence of the notion of sustainable development has led to a shift from single-criterion-oriented evaluations to more integrated assessments covering the different dimensions of livestock performance: economics, social, health, welfare and the environment. Many methodological questions arise from the confrontation of these dimensions which have different measurement scales. How to support the evolution of livestock systems with a multicriteria assessment? How to well define the purposes of the evaluation? How to manage the multiplicity of information? How to interpret a set of indicators together when they take opposite values? Which tools to mobilize? Should we use or create our own assessment tools? Livestock farming with its multiple stakes at different scales herd, farm and territory, is particularly concerned with these issues. The participation of the actors in the evaluation process is also a question to be decided: are they co-designers of the method, validators, surveyors or only recipients of the results? These different elements will lead to the definition of the best methods of mathematical data processing. This data processing is intimately linked to how the results are represented and how they are used. This paper discusses the different uses of multicriteria evaluations applied to livestock systems and the key methodological points.

LAIREZ J., FESCHET P., BOTREAU R., BOCKSTALLER C., FORTUN-LAMOTHE L., BOUVAREL I., AUBIN J., 2017. L'évaluation multicritère des systèmes d'élevage pour accompagner leurs évolutions : démarches, enjeux et questions soulevées. *INRA Prod. Anim.*, 30, 255-268.