

Evaluation agri-environnementale et choix des indicateurs : acquis, enjeux et pistes

**Bockstaller C.¹, Cariolle M.², Galan M-B. ³, Guichard L.⁴, Leclercq C.⁵, Morin A.⁶,
Surleau-Chambenoit C.⁶**

¹ UMR Agronomie Environnement, INRA-Université de Nancy-Lorraine, 54505 Vandoeuvre les Nancy

² ITB, 45 rue de Naples, 75008 Paris

³ AMBRE Développement, 19 allée des Goyaviers, Belair Desrozières, 91170 Petit-Bourg.

⁴ UMR Agronomie, INRA-AgroParisTech, 78850 Thiverval-Grignon

⁵ Institut Polytechnique Lasalle Beauvais, 19 Rue Pierre Waguët, 60000 Beauvais

⁶ Agro-Transfert Ressources et Territoires, 2 Chaussée de Brunehaut, 80200 Estrées-Mons

Correspondance : christian.bockstaller@colmar.inra.fr

Résumé

Depuis les années 90, la prise de conscience environnementale et l'émergence du développement durable a conduit à une demande croissante de moyens opérationnels d'évaluation agri-environnementale sous forme d'indicateurs. En a suivi une multiplication et un foisonnement d'initiatives proposant des indicateurs et des méthodes d'évaluation basées sur une liste plus ou moins organisée d'indicateurs. Face à cette offre très large, l'utilisateur potentiel peut rester perplexe et s'y perdre. Cela l'amène souvent à un choix contingent à la méthode dont il a entendu parler, disponible dans son environnement professionnel proche, sans qu'il s'interroge nécessairement sur l'adéquation de son choix à ses besoins et questionnements. Cet article présentera successivement les étapes clés dans une démarche de choix d'indicateurs, des éléments sur l'offre disponible pour « dégrossir » le choix, en abordant notamment la question de l'analyse de cycle de vie et de différents types d'indicateurs, et les aides disponibles pour réaliser concrètement un choix. Dans une dernière partie, nous discuterons des questions liées à l'évaluation agri-environnementale à l'aide d'indicateurs, nécessitant un effort de recherche supplémentaire.

Mots-clés : qualité prédictive, agrégation, changement d'échelle, analyse de cycle de vie

Abstract: Agri-environment evaluation and choice of indicators: achievements, stakes and possibilities

In the 90s, the growing concern for environmental issues and the emergence of the concept of sustainable development led to an increasing demand for operational assessment means as indicators. This led to an "explosion" of indicators and assessment methods based on a list of indicators, more or less organized. The potential user facing this very broad offer can be lost. This often leads him to a contingent selection of a method which he heard about or which is available in his professional context, without asking whether the selected method meets his needs. This article will successively present the key points in the selection procedure of an indicator, some insights in the available offer of indicators, tackling among others the life cycle analysis approach, and the existing help available to carry out practically the choice. In a last section, we will discuss the further needs of research on environmental assessment by means of indicators.

Keywords: predictive quality, aggregation, upscaling, life cycle analysis

Introduction

La question des indicateurs agri-environnementaux a émergé dans les années 1990 avec l'apparition sur le devant de la scène de la problématique environnementale et plus généralement du débat autour de la durabilité suite à la conférence de Rio en 1992. La mise en évidence des effets négatifs collatéraux ou secondaires de l'intensification de l'agriculture et plus généralement des politiques de croissance, voire de certaines innovations technologiques a conduit les acteurs à porter une attention de plus en plus forte à l'étape de l'évaluation, exprimant un besoin accru de méthodes d'évaluation et de suivi. Celle-ci est devenue aujourd'hui indispensable en décision publique, dans la recherche, dans la conception de solution innovante, dans les démarches de progrès comme dans la norme ISO 14000, dans le travail des ONG, etc. (López-Ridaura *et al.*, 2005 ; Niemeijer et de Groot, 2008). D'une manière générale, le recours à des indicateurs à la place d'une mesure directe des impacts s'explique aisément par les problèmes de faisabilité d'une approche directe (coût de la mesure, temps d'acquisition des réponses,...), des problèmes de métrologie face à la complexité des concepts, systèmes ou processus (Gras *et al.*, 1989 ; Maurizi et Verrel, 2002). Ainsi, ce besoin de procédure d'évaluation et les difficultés de mise en œuvre de mesures directes en routine à grande échelle et dans les conditions de la pratique, en dehors du champ de l'expérimentation scientifique, expliquent ce formidable développement de travaux sur les indicateurs, notamment dans le domaine agricole (Rosnoblet *et al.*, 2006 ; Bockstaller *et al.*, 2009).

Face à ce foisonnement d'indicateurs et de méthodes d'évaluation, basées sur des listes plus ou moins organisées d'indicateurs, l'utilisateur potentiel peut cependant rester perplexe et s'y perdre. Cela l'amène souvent à un choix contingent à la méthode dont il a entendu parler, disponible dans son environnement professionnel proche, sans qu'il s'interroge sur l'adéquation de son choix à ses besoins et questionnements. Pour pallier le risque de s'engager dans une mise en œuvre, parfois lourde et coûteuse, qui ne se révélerait pas *in fine* satisfaisante, il est devenu indispensable de fournir des éléments de choix aux utilisateurs potentiels. Cet article présentera successivement, les étapes clés dans une démarche de choix d'indicateurs, des éléments sur l'offre disponible pour « dégrossir » le choix, et les aides disponibles pour réaliser concrètement un choix. Dans une dernière partie, nous discuterons des questions liées à l'évaluation agri-environnementale, nécessitant un effort de recherche supplémentaire.

1. Les étapes clés du choix d'un indicateur ou d'une méthode d'évaluation

1.1 Des questions préalables à éclaircir

Cette étape est indispensable pour bien choisir la méthode, en fixer en quelques sortes le cahier de charges. La clarification de questions préalables permettra d'éviter de « foncer tête baissée dans le choix des indicateurs », ce qui entraîne souvent des choix contingents liés à la connaissance de telle ou telle méthode.

1.1.1 Nécessité d'un diagnostic initial (pourquoi évaluer ?)

Il s'agit de dresser un état des lieux de la situation, l'état du système étudié, quels sont les enjeux environnementaux pour lesquels il y a problème, les points à améliorer, etc., mais aussi les facteurs explicatifs, quelles sont les pratiques agricoles impliquées, les sous-parties du système impliquées (Ex : parcelles à risque), etc. Cette partie peut s'appuyer sur des études préalables, sur une expertise locale et doit être impérativement réalisée pour accorder les différents intervenants et bénéficiaires de l'évaluation et pour éviter les débats stériles et récurrents sur les responsabilités (Ex : part des activités agricoles/activités non agricoles dans le déterminisme des impacts).

1.1.2 Préciser la finalité et les utilisateurs (pour quoi et pour qui ?)

Les finalités dépendent naturellement des « utilisateurs », terme qui peut porter sur les commanditaires de l'évaluation, bénéficiaires des résultats, qui ne sont pas nécessairement les opérateurs. Cela peut être des décideurs de l'Etat, de collectivités ou les gestionnaires, des agronomes de la R&D ou des conseillers travaillant directement avec les agriculteurs, voire des agriculteurs ou groupes d'agriculteurs, et d'autres acteurs (protection de la nature, associations de consommateurs).

Dans le cadre agri-environnemental, la plateforme Plage notamment (Surleau-Chambenoit *et al.*, 2013)¹ a identifié six finalités, issues de l'étude des usages d'un panel de méthodes d'évaluation (Leclercq *et al.*, 2011) :

- acquérir des références sur des systèmes de production
- conseiller et accompagner le changement
- faire des simulations de changement de pratiques
- suivre les pratiques agricoles (tableau de bord)
- réaliser un management environnemental (en vue d'une certification ISO 14001)
- vérifier la conformité vis-à-vis de la réglementation

Il est à noter que cette dernière finalité portant sur la vérification de la conformité réglementaire qui suscite une demande des administrations, est un cas particulier à différencier en réalité de l'évaluation proprement dite (Chémery et Boissier, 2002). Dans ce contexte, il est nécessaire d'utiliser une information vérifiable, sinon il peut exister un risque de fraude (Oenema *et al.*, 2003).

D'une manière générale, l'évaluation peut concerner des situations réelles *a posteriori* (*ex post*), soit en cours d'action ou à la fin, comme elle peut porter sur des situations virtuelles *a priori* (*ex ante*), tels un ensemble de solutions, d'actions potentielles, ou un scénario, pour répondre à une question de type « que se passerait-t-il si je change telle pratique ? ». La finalité peut être d'avoir une information sur l'état, les performances du système, etc, ce qui peut être suffisant pour certains utilisateurs. Pour d'autres usages, il sera nécessaire de pouvoir identifier à partir de l'évaluation les éléments à améliorer, les pratiques à changer, les solutions à introduire; dans ce cas l'évaluation doit conduire à une aide à la décision.

Enfin, une autre finalité souvent évoquée est de communiquer sur les résultats de l'évaluation (Girardin *et al.*, 2005).

1.1.3 Définir l'objet de l'évaluation (on évalue quoi ?) : le choix des enjeux environnementaux

L'évaluation peut porter sur les différentes étapes d'un programme d'action : la pertinence des objectifs vis-à-vis des enjeux identifiés, la cohérence des moyens (humains, financiers, méthodes, leur répartition spatio-temporelle, les changements de pratiques, les impacts proprement dits sur l'environnement (résultats) comme l'efficacité (résultats/moyens) (CORPEN, 2006). Ces différents objets intéresseront différents acteurs.

L'évaluation des résultats, des effets ou des performances environnementales, qui pose le plus de questions scientifiques et qui nous intéressera par la suite amène à un second choix, celui des enjeux liés aux différents compartiments et composantes environnementales. Il est à noter que certains auteurs parlent d'impacts (analyse de cycle de vie, van der Werf *et al.* 2011), de compartiments environnementaux (Girardin *et al.*, 2005), plus récemment de services écosystémiques, ou, de manière plus théorique, de propriétés systémiques (López-Ridaura *et al.*, 2005).

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

Plutôt que de dresser une liste d'enjeux qui déterminera le choix des indicateurs, il est possible de les hiérarchiser en fonction du système étudié et de l'échelle à laquelle tel ou tel impact est déterminé. Dans la Figure 1, nous proposons de différencier :

- Les enjeux directs pour le système, tel le sol de la parcelle étudié, ou encore la qualité de l'eau en sortie de bassin versant (élémentaire)
- les contributions à des enjeux locaux (qualité de l'eau de la nappe, ou du cours d'eau en aval, alimenté par le bassin versant)
- les contributions aux enjeux globaux, tel le changement climatique

Les deux premiers cas sont en fait locaux. Dans la perspective de la durabilité, il est important de ne pas se focaliser sur les enjeux directs, voire la contribution aux enjeux locaux mais aussi de prendre en compte la dimension globale pour éviter qu'une solution pour répondre aux enjeux directs ne puisse se révéler préjudiciable par des effets collatéraux pour les enjeux globaux. Par exemple, la réduction des herbicides par des moyens mécaniques peut entraîner une augmentation de la consommation énergétique (Leclerc *et al.* dans ce numéro). Le terme de contribution implique que le système étudié n'entre pas seul dans le déterminisme de l'impact lié à un enjeu. Il est à noter aussi que cette classification des enjeux est fonction de la définition du périmètre et des échelles du système.

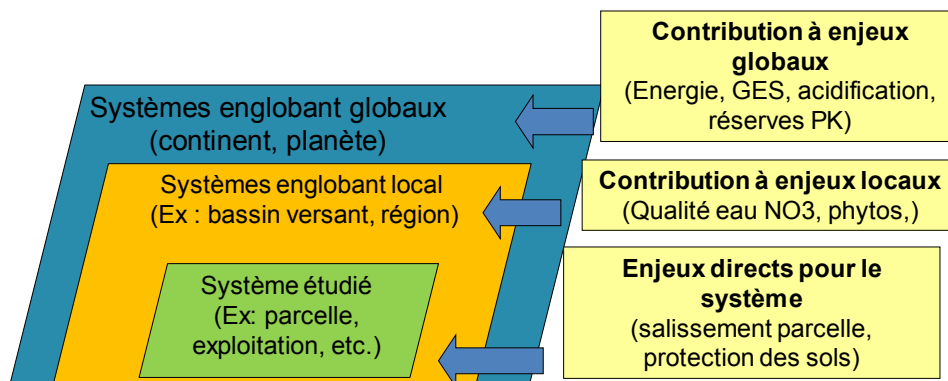


Figure 1 : Une classification des enjeux en fonction du système et des échelles spatiales.

1.1.4 Définir les limites du système évalué et les échelles de l'évaluation (spatiales où ? temporelles quand ?)

La définition des échelles de calcul au niveau spatial comme temporel passe par la définition de deux notions importantes : (1) la *résolution*, échelle à laquelle on veut avoir les résultats de base et qui implique la saisie à cette échelle et (2) l'*étendue* qui donne les limites de l'espace ou de la période étudiés. Les acteurs peuvent être intéressés par une méthode fournissant par exemple des résultats uniquement au niveau de l'exploitation et de l'année (résolution et étendue identiques) ou encore avoir les résultats parcellaires et mensuels (résolution) au niveau de l'exploitation et pour une année (étendue).

La définition du système ne va pas directement de pair avec la définition de l'échelle. En effet, évaluer les impacts environnementaux d'une culture à la parcelle peut se limiter aux impacts directs pour les différents enjeux sans tenir compte de l'amont (impacts liés à la production des intrants) ou l'aval (impacts liés à l'utilisation de la production, à ses déchets). Cette question de la définition du périmètre du système est un passage obligé dans les méthodes basées sur l'analyse de cycle de vie (van der Werf *et al.*, 2011) et devrait être abordée explicitement dans tous les cas (*cf.* paragraphe 2.2).

1.1.5 Préciser qui réalise l'évaluation et quelles sont les contraintes de l'évaluation (budget, temps, données disponibles, comment ?)

A côté des considérations théoriques, le choix des indicateurs et des méthodes sera fonction des moyens dont dispose l'utilisateur, son budget et le temps disponible pour accéder aux données, mettre en œuvre les calculs et les restituer, les données disponibles, etc.

1.2 Choix des unités fonctionnelles et des valeurs de référence

Bien que dans de nombreux cas, les unités fonctionnelles des indicateurs proposées par les différentes méthodes soient imposées (et souvent en ha), l'unité doit être choisie en analyse de cycle de vie par rapport à la fonction attribuée au système. En l'occurrence pour l'agriculture, il est possible d'exprimer les indicateurs par unité économique si la fonction est d'assurer la rentabilité, par unité de production (exemple : kg de céréales produits, litre de lait produit, etc...), si l'accent est mis sur les volumes de production, ou encore par unité de surface si la fonction est environnementale. Ce choix n'est pas neutre du tout et peut changer le classement entre systèmes intensifs et extensifs. On voit que dans l'exemple suivant, choisir une unité de production privilégie les systèmes intensifs (impacts globaux et usage des terres à production équivalente moindre), et celui d'une unité de surface favorise les systèmes plus extensifs, ici en agriculture biologique (Tableau 1). S'il est possible de choisir l'unité, il paraît pertinent d'utiliser pour les impacts globaux non liés à une zone donnée, une unité de production et pour les impacts locaux liés à une zone définie (qui ne peuvent donc être compensés) une unité de surface (Halberg *et al.*, 2005). Ce choix des unités fonctionnelles est donc à revoir dans de nombreux cas pour éviter le piège du choix automatique de l'unité surfacique.

Tableau 1 : Effet du choix d'une unité fonctionnelle sur les performances de systèmes porcins en fonction de leur intensification (exemple tiré de Halberg *et al.*, 2005).

	Indicateur/kg		Indicateur/ha	
	Intensif	Agribio	Intensif	Agribio
Impact local (eutrophisation kg PO ₄ -eq)	100	104	100	57
Impact global (effet de serre (kg CO ₂ -eq))	100	173	100	95
Utilisation des terres (m ² /an)	100	182	100	100
Production (kg porc)	100	100	100	55

Le choix d'une valeur de référence est indispensable pour donner du sens à un indicateur et le différencier de la valeur brute d'une variable (Girardin *et al.*, 2005). Cette référence peut être absolue sous forme de seuil, de valeur cible ou encore relative, via une valeur d'un système moyen, ou représentatif, ou d'une année de référence comme l'année initiale d'une action. Dans le domaine environnemental, des valeurs de référence absolue seraient à souhaiter mais manquent souvent. Leur fixation ne relève pas juste de choix scientifique mais aussi politique. Dans tous les cas, il conviendra de se méfier de référence implicite comme par exemple une valeur de zéro pour le solde azoté à l'équilibre (Bockstaller *et al.*, 2012).

1.3 Choisir en connaissant la sensibilité et la qualité prédictive des indicateurs

Dans une liste d'indicateurs qui répondent aux questions préalables, tous ne présentent pas les mêmes qualités scientifiques en termes de sensibilité, i.e. capacité à discriminer des systèmes a priori jugés différents. Par ailleurs, si les indicateurs doivent fournir une information sur l'effet des pratiques agricoles sur l'environnement, il est important de disposer de l'information sur la qualité prédictive de l'indicateur, de sa capacité de rendre compte des effets. Bien que des auteurs aient proposé des tests, cette information reste encore très éparse et peu de travaux ont visé à la caractériser de manière systématique. Un indicateur aussi utilisé que le solde azoté présente par exemple une faible capacité prédictive de la pollution au nitrate s'il est calculé annuellement à l'exception de systèmes très excédentaires (Bockstaller *et al.*, 2012). Des données ont été publiées sur des indicateurs simples de biodiversité (Dennis *et al.*, 2009). Ils portent généralement sur des comparaisons, valeur de l'indicateur-diversité d'un groupe taxonomique, ceci pour une gamme très différenciée de situations (Ex : paysages de bocage et prairies vs. paysages ouverts céréaliers), et non juste pour des situations proches (Ex : des paysages céréaliers présentant différents niveaux d'intensification et d'ouverture). L'indicateur I-Phy de la méthode INDIGO a fait l'objet de travaux avec trois autres indicateurs dans des comparaisons avec des teneurs en substances actives en sortie de bassins viticoles (Girardin *et al.*, 2007).

A défaut d'avoir de telles informations, une première étape peut consister à resituer l'indicateur sur la chaîne causale ou réseau causal (Niemeijer et de Groot, 2008) et voir dans quel mesure il peut être associé à des processus ou les intégrer.

2. Le choix des indicateurs et des méthodes

2.1 Analyse de Cycle de Vie vs. méthodes « indicateurs »

Durant ces dernières années, un certain nombre de travaux d'évaluation ont reposé sur l'analyse de cycle de vie (ACV) qui vient de l'industrie (van der Werf *et al.*, 2011). Cette méthode qui fait l'objet d'une normalisation ISO repose sur une démarche rigoureuse, évaluant les impacts tout au long du cycle de production et non seulement les impacts directs du système, ceci pour éviter les reports de pollution au long de la chaîne de production.

D'autres méthodes couramment utilisées en France, telles IDEA, INDIGO, le DAE-G (Bockstaller *et al.*, 2008a) reposent surtout sur des indicateurs évaluant les effets directs au champ ou sur l'exploitation. Cependant, pour l'enjeu énergie, ces méthodes incluent un calcul de la consommation indirecte venant de la fabrication des intrants. Par ailleurs, des algorithmes de calcul utilisés par ces méthodes pour évaluer les impacts directs sont utilisés en ACV, tel le solde azoté (Bockstaller *et al.*, 2012) ou pourraient l'être.

Ainsi, il ne s'agit pas d'opposer l'ACV et le reste des méthodes d'évaluation. Vu la lourdeur de l'ACV pour la mise en œuvre mais aussi l'interprétation des résultats, et les questions qui restent à trancher (Bockstaller *et al.*, 2012), l'ACV ne doit pas être présentée comme la solution clé en main et universelle pour toute situation d'évaluation. Elle nous semble pertinente pour comparer des systèmes très différents (Ex : systèmes herbagers vs. système maïs-soja, ou des filières). Mais pour des systèmes assez proches, comme différents systèmes céréaliers, une ACV systématique ne paraît pas forcément nécessaire, les impacts indirects (en amont du champ) ne pouvant varier que faiblement. Une confrontation des sorties entre une méthode évaluant les effets directs et l'ACV est prévue dans le projet SCEP-DEPHY de l'appel d'offre PSPE 2012 d'Ecophyto.

2.2 Quels types d'indicateurs pour quels enjeux ?

Deux typologies pression/état/réponse de l'OCDE et une dérivée, force-motrice/pression/état/impact/réponse (DPISR) de l'Agence Européenne pour l'Environnement (EEA), inspirées de la chaîne causale, présentée dans la Figure 2 sont utilisées couramment au niveau national et international (Maurizi et Verrel, 2002). Cependant, ces typologies donnent l'impression d'une linéarité et de processus mono-factoriels, ce qui n'est généralement pas le cas (Niemeijer et de Groot, 2008). Par ailleurs, des notions comme la pression sont ambiguës. Certes, la pression fait référence aux causes, mais selon les auteurs, elle est reliée aux pratiques agricoles (CORPEN, 2006) ou encore aux émissions (Bockstaller *et al.*, 2012), avec pour conséquence un choix d'indicateurs totalement différents. Les émissions peuvent aussi être considérées comme des « effets » de premier niveau sur la chaîne causale, d'où le terme d'indicateur d'effets utilisé par van der Werf et Petit (2002). Ces indicateurs d'effet ne peuvent être comparés aux indicateurs de moyens basés sur les pratiques agricoles. La même chose est vraie pour les indicateurs d'état qui englobent, dans la typologie OCDE, l'état et l'impact de la typologie de l'EEA. Ceci explique aussi les difficultés de classification de ces indicateurs dans ces typologies que certains ont rencontrées dans plusieurs groupes de travail (Girardin *et al.*, 2005).

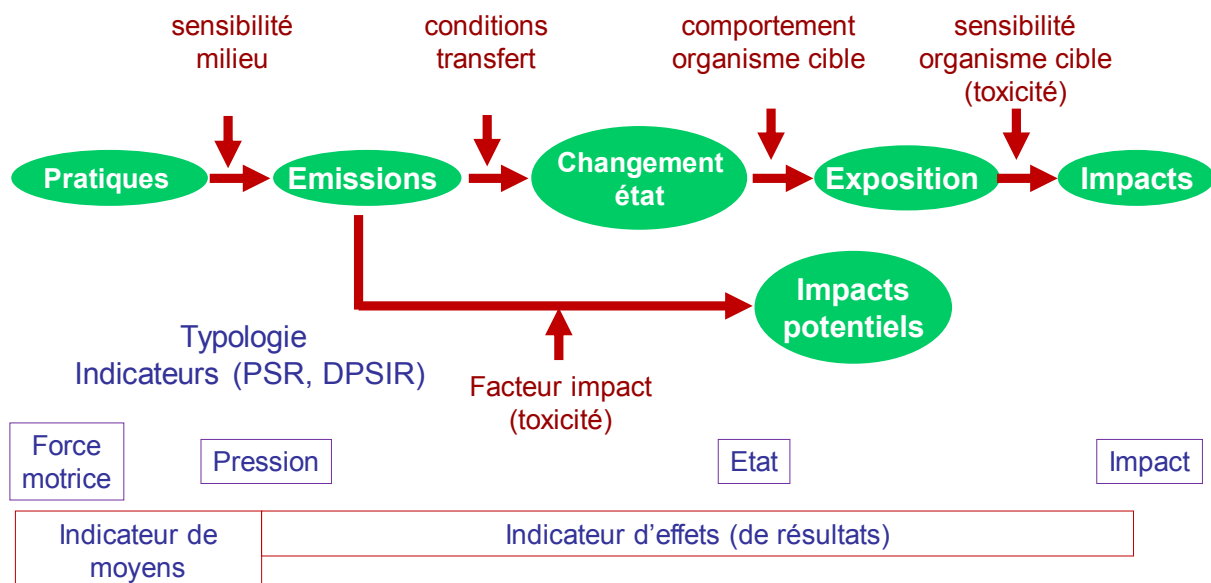


Figure 2 : Typologie des indicateurs en fonction des modalités d'obtention des valeurs et de leurs atouts et limites (inspirée de Bockstaller *et al.*, 2012))

A partir de ce qui vient d'être dit sur les limites des typologies utilisées, nous préconisons une typologie plus claire fondée sur la chaîne causale en partant d'indicateurs de pratiques agricoles, d'émissions, de changement d'état, d'impacts, etc. Nous proposons de croiser cette première typologie par une seconde qui fait référence à la nature de la grandeur derrière l'indicateur (Figure 3) :

- Les indicateurs *simples* qui sont basés sur une variable ou une combinaison mathématique simple de variables de pratiques souvent, sous forme de ratio ou de solde. Le terme « simple » s'applique à la structure de l'indicateur, sachant que leur mise en œuvre à l'échelle d'un pays comme la France peut poser de nombreux problèmes opérationnels. Ils sont cependant en général plus faciles à utiliser que les suivants mais intègrent faiblement les processus et pris isolément ne donnent qu'une estimation incertaine des processus et des impacts (CORPEN, 2006), n'étant pas prévus pour cela.

- A l'autre extrême, un groupe des indicateurs reposant sur des *mesures* de terrain (Ex : des émissions de nitrate jusqu'aux mesures d'impact sur une population d'organismes vivants). Ceux-ci, si la métrologie est maîtrisée, intègrent les processus et peuvent donner une image assez exacte de la situation mais sont généralement lourds (et coûteux) à mettre en œuvre. Ils ne permettent pas non plus de tracer directement les causes, ce qui nécessite des mesures complémentaires.
- Les indicateurs *prédictifs* reposant sur une fonction estimatrice, modèles opérationnels à nombre réduit de variables accessibles ou modèles complexes à base mécanistique. Ils présentent un compromis entre les deux types précédents et surtout permettent de relier la variable estimée à des variables d'entrées représentant des causes.

Ces groupes d'indicateurs ne sont pas opposés entre eux mais peuvent être complémentaires en fonction des finalités et des utilisateurs : les premiers sont utiles pour suivre les changements de pratiques, les seconds permettront d'acquérir des références sur les impacts des systèmes tandis que les indicateurs prédictifs sont utiles pour le conseil, pour réaliser des simulations en évaluation *ex ante*, par exemple de systèmes de culture innovants.

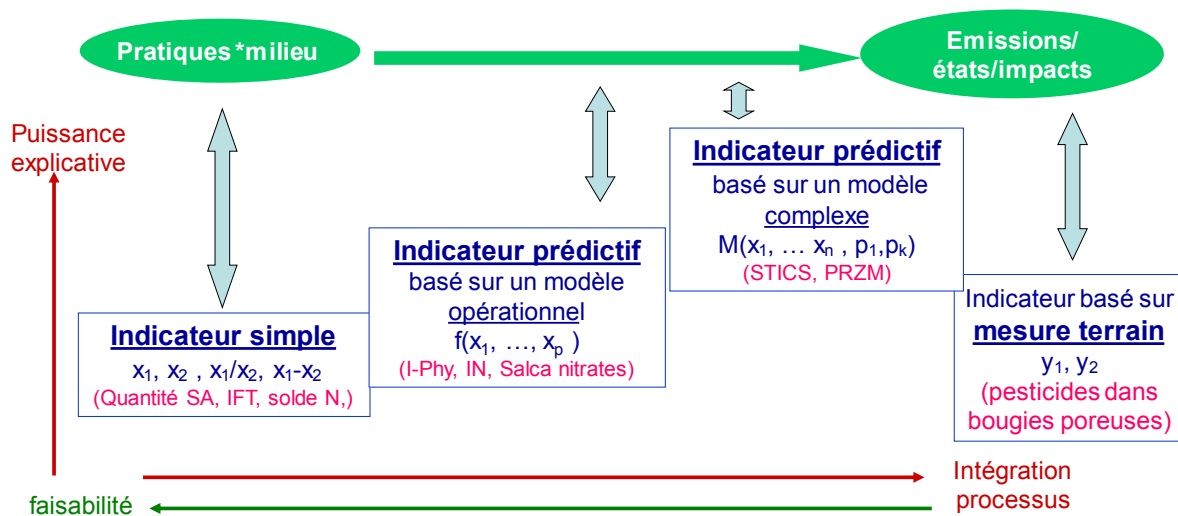


Figure 3 : Typologie des indicateurs en fonction des modalités d'obtention des valeurs et de leurs atouts et limites (inspirée de Bockstaller *et al.*, 2012)

2.3 La question de l'agrégation multi-enjeu

Si on excepte des indicateurs simples basés sur une variable, la question de l'agrégation est très présente dans les indicateurs. Les indicateurs prédictifs reposent sur des modèles qui sont construits selon des fonctions mathématiques ou des arbres de décisions intégrant des connaissances sur les processus ou des dires d'experts. La question de l'agrégation devient plus délicate quand il s'agit de passer à un niveau supérieur en reprenant la chaîne causale de la Figure 2 et que les indicateurs ne sont pas exprimés dans la même unité. Un premier niveau est l'agrégation, pour un même enjeu, de plusieurs types d'émissions impliqués dans un impact, tels les différents gaz à effet de serre. Les travaux en ACV ont apporté des solutions intéressantes en passant par des facteurs d'impacts qui convertissent toutes les émissions en une unité, par exemple en équivalent CO₂ pour les gaz à effet de serre. Mais dans l'évaluation de la durabilité, se pose la question de l'agrégation multicritère ou multi-enjeu pour arriver à un indicateur *composite* (Girardin *et al.*, 2005).

Certains auteurs s'y refusent en raison des pièges de l'agrégation, tels « l'addition des choux et des carottes », la subjectivité de la pondération ou encore la perte d'information liée à l'agrégation par rapport aux indicateurs non agrégés. Cependant, nous constatons que dans beaucoup de cas, les acteurs ont besoin à un moment ou un autre de la démarche d'évaluation, d'un indicateur agrégé pour fournir une information synthétique qui conclue, ou pour comparer, sélectionner des actions, des systèmes. L'évaluation détaillée indicateur par indicateur pour les différents enjeux reste indispensable pour identifier les points forts et faibles dans le cas d'un travail de conseil, ou pour expliquer la valeur agrégée. C'est bien ces allers-retours que nous préconisons.

Différentes méthodes existent pour une agrégation « composite » pertinente (Bockstaller *et al.*, 2012) :

- Des méthodes reposent sur la normalisation des indicateurs par une approche monétaire ou celle, physique, de l'empreinte écologique qui posent cependant de nombreuses questions méthodologiques, comme l'évaluation des biens non marchands pour la première ou la définition des facteurs de pondération pour la seconde.
- A l'autre extrême, des méthodes d'analyse multicritère basées sur le surclassement, comparant deux à deux les critères ou indicateurs ne proposent pas un indicateur unique mais un classement ou un tri des actions et systèmes étudiés. Elles ne sont pas adaptées à l'utilisation d'information mixte, qualitative et quantitative et utilisent généralement un nombre limité de critères.
- Les approches « mixtes » reposant sur des arbres de décisions qui peuvent s'emboîter, et combinant l'information sous forme de règles « si alors ». Elles ont l'avantage de s'exprimer dans un langage courant. Ces dernières années, un outil informatique DEXi a grandement facilité l'accès à cette approche et a permis le développement de méthodes multicritères évaluant par exemple la durabilité de systèmes de culture, MASC (Craheix *et al.*, 2012 ; Colomb et Bergez, 2013) ou DEXiPM pour les systèmes de protection des cultures (Messéan *et al.*, 2010). Ces approches utilisent pour MASC des indicateurs quantitatifs tels ceux d'INDIGO mis en classe et des indicateurs qualitatifs, ou pour DEXiPM des indicateurs qualitatifs simples de pratiques et des variables de contexte.

Les approches « mixtes » peuvent présenter deux inconvénients : le problème des valeurs de limites de classe qui peut être résolue par l'utilisation de sous-ensembles flous et la subjectivité des règles de décision. La méthode SIRIS reposant sur un système de pénalisation qui a été utilisée pour agréger des variables liées au transfert des produits phytopharmaceutiques limite cette subjectivité (Pierre et Colmar, 2009). Dans le cadre du projet Casdar INDIBIO (2011-2013) dédié aux indicateurs de biodiversité, nous sommes en train de développer une méthode d'agrégation combinant les avantages de ces différentes approches.

3. Quelle aide au choix pour l'utilisateur potentiel ?

Ce contexte, depuis la fin des années 90, de multiplication et de foisonnement d'indicateurs, et de méthodes d'indicateurs, rend indispensable la fourniture d'une aide aux utilisateurs potentiels.

Les premières initiatives allant dans ce sens ont consisté en des synthèses descriptives des méthodes existantes (van der Werf et Petit, 2002 ; Bockstaller *et al.*, 2008a) ou pour un enjeu donné comme l'azote (CORPEN, 2006 ; Bockstaller *et al.*, 2012), les produits phytopharmaceutiques (Devillers *et al.*, 2005) ou encore la biodiversité (Bockstaller *et al.*, 2011). Dans ces travaux, les indicateurs ont été analysés, comparés à l'aide de critères descriptifs, ou encore décrits dans des fiches (Devillers *et al.*, 2005 ; CORPEN, 2006). Dans la synthèse du CORPEN (2006), les fiches contiennent notamment des conseils d'interprétation.

Plus récemment, un groupe de travail porté par le GIS Grande Culture à Hautes Performances Economique et Environnementale (GIS GC-HP2E) est en train de développer la base de données INDIC fournissant une information descriptive de base (enjeux, mode de calcul, etc.) sur 2325 indicateurs (1468 environnementaux, 485 économiques, 376 sociaux) issues de 112 méthodes identifiées dans la littérature scientifique et grise, et internet à l'échelle mondiale.

Les travaux cités ci-dessus ne présentent que peu ou pas de critères informant l'utilisateur sur les qualités, atouts et limites des méthodes pour l'aider à sélectionner celle qui répond le mieux à ses besoins. La synthèse du CORPEN (2006) fournit deux critères : « la pertinence » et « la faisabilité » pour les indicateurs à l'échelle de la parcelle tandis que Devillers *et al.* (2005) en proposent quatre : « la lisibilité », « la faisabilité », « la reproductibilité » et « la pertinence pour les utilisateurs ». Aucune explication précise n'est donnée à l'attribution des valeurs pour ces critères. Dans le projet transfrontalier COMETE, Bockstaller *et al.* (2006) sont allés plus loin en développant une grille d'analyse de 15 critères regroupés en trois volets « pertinence scientifique », « faisabilité » et « utilité ». Pour chaque critère, noté entre 1 et 5, des règles d'attribution ont été proposées. Ceci a permis de comparer quatre méthodes portant sur les échelles de l'exploitation agricole et parcellaire pour les trois premières, INDIGO SALCA (méthode suisse d'analyse de cycle de vie) et deux méthodes allemandes REPRO et KUL. D'autres critères ont été utilisés par d'autres auteurs et recensés par Niemeijer et de Groot (2008).

Cependant, tous ces travaux « sur papier » n'offrent pas d'interaction directe avec l'utilisateur. C'est ce besoin auquel ont voulu répondre les partenaires de la plateforme PLAGE¹ en offrant une aide au choix interactive pour laquelle l'utilisateur potentiel à la recherche d'une méthode saisit un certain nombre d'informations pour identifier des méthodes qui correspondent à ses besoins, sur les préalables à l'évaluation (finalités, enjeux, échelles, etc.), puis dans un second temps sur des critères plus précis de faisabilité notamment (Surleau-Chambenoit *et al.*, 2013). A notre connaissance, il n'existe pas d'autre initiative de ce type. Nous n'avons trouvé aucun article scientifique sur le Web of Science et une recherche internet sur Google n'a révélé que deux sites allant dans ce sens : celui de l'Agence européenne pour la sécurité et la santé au travail (EU-OSHA)² qui permet de rechercher de diverses manières les outils d'évaluation des risques, notamment par sujet, par secteur ou par pays, et le site GUIDE³. Celui-ci a été développé dans un projet lancé par l'ONEMA et l'INRA sur le choix des indicateurs traitant des produits phytopharmaceutiques et de la qualité de l'eau qui a repris le travail de Devillers *et al.* (2005) pour cet enjeu (Keichinger *et al.*, 2013). Certains partenaires ont aussi participé au projet Plage¹.

4. Discussion : quels besoins de recherche ?

Malgré une offre très abondante en matière d'indicateurs agri-environnementaux ou de méthodes multicritères, le choix peut devenir beaucoup plus restreint pour certains enjeux. Par exemple, bien qu'il existe de nombreux indicateurs traitant de la biodiversité, l'offre est bien plus pauvre quant aux indicateurs de type prédictif (Figure 3) (Bockstaller *et al.*, 2011). Ceci est encore plus vrai pour les risques de compaction du sol ou de pollution par les métaux lourds et autres polluants, les risques liés à la gestion des déchets ou encore les bruits et les odeurs. De même, des nouveaux polluants des milieux tels les antibiotiques manquent tout naturellement d'indicateurs prédictifs en raison de manques de connaissances sur les processus.

² https://osha.europa.eu/fr/practical-solutions/risk-assessment-tools/index_html

³ <http://www.myrmesys.com/guide/>

Si l'offre est abondante à l'échelle de la parcelle, de l'exploitation pour différents types d'indicateurs et à un degré moindre aux échelles plus grandes (bassin versant, région, pays, etc.) pour les indicateurs de type simple ou basés sur des mesures, la question du changement d'échelle continue d'être d'actualité.

Aux grandes échelles, la validité statistique des indicateurs de pratiques et ceux basés sur des mesures est constamment un souci car il n'est pas possible d'avoir accès à une information exhaustive. A cette échelle, l'accès aux données reste encore un problème notamment pour les pratiques agricoles.

Par ailleurs, l'intégration des zones non agricoles dans l'évaluation, leur rôle de compensation ou de source de pollution exogène (CORPEN, 2006) sont à prendre en compte. Cette question se pose en particulier lors du passage du calcul d'indicateurs parcellaires à des échelles plus grandes. Une méthode comme TERRIT'EAU (Gascuel *et al.*, dans ce numéro) est un exemple de travail à cette échelle comme l'adaptation de l'indicateur I-Phy d'INDIGO à l'échelle du bassin versant en prenant en compte la connectivité des parcelles au réseau hydrographique (Wohlfahrt *et al.*, 2010).

Le travail de caractérisation et d'évaluation de cette offre abondante est bien entamé avec très récemment les travaux de PLAGE (Surleau-Chambenoit *et al.*, 2013) pour des méthodes d'évaluation, de GUIDE (Keichinger *et al.*, 2013)³ pour les indicateurs traitant de l'enjeu produits phytopharmaceutiques et de qualité de l'eau. Ceux-ci fournissent une réelle aide au choix, interactive, à la différence des synthèses papier comme celle du CORPEN sur indicateur azote. Cependant, ces sites peuvent renvoyer à un ou plusieurs indicateurs ou méthodes qui sont décrites à l'aide d'une fiche proche des synthèses papiers. Certaines synthèses présentent une information supplémentaire et importante qui n'est pas à négliger avant toute mise en œuvre d'un indicateur ou d'une méthode. Elle peut porter sur les conseils à l'interprétation en pointant les limites (CORPEN, 2006). Y manque aussi encore trop souvent une information suffisante sur la question des valeurs de références qui mériterait un investissement supplémentaire (Bockstaller *et al.*, 2008b). Pour les indicateurs environnementaux, cela passe aussi par une connaissance des relations entre valeur de l'indicateur et données mesurées. Bien qu'il existe des méthodes de comparaison (Bockstaller *et al.*, 2008b), une évaluation systématique de la qualité prédictive des indicateurs pour un enjeu donné fait défaut, ceci faute en grande partie de jeux de données suffisants. C'est l'ambition du projet EQUIPE qui vient de débiter dans le cadre de l'appel d'offre PSPE 2012 d'Ecophyto, de regrouper plusieurs jeux de données et d'évaluer un ensemble de différents indicateurs traitant de l'enjeu « produits phytopharmaceutiques et qualité de l'eau ». Une question connexe est celle de l'incertitude autour du calcul ou d'une mesure d'un indicateur. Celle-ci est bien documentée sur certains indicateurs comme les soldes azotés ou l'ACV (Bockstaller *et al.*, 2012) et serait à étendre à d'autres méthodes. Un travail a été réalisé pour l'indicateur I-Phy d'INDIGO en prenant en compte la variabilité des valeurs des propriétés des substances actives (Assaghir, 2010). Enfin, une comparaison des sorties d'indicateurs afin de voir dans quelle mesure ils conduisent à des préconisations identiques serait à répéter à partir de la méthode proposée dans le projet COMETE (Bockstaller *et al.*, 2006) ou celle utilisée par Agro-Transfert Ressources et Territoires (Galan *et al.*, 2007).

Dans la démarche de validation avec au cœur l'évaluation de la qualité prédictive, Bockstaller *et al.* (2008b) ont proposé une étape de validation par les utilisateurs finaux. Il s'agit donc d'évaluer dans quelle mesure un indicateur est utilisé, et comment il répond aux besoins des utilisateurs. Cette étape est importante pour améliorer et adapter l'offre à la demande. Dans le cas de MASC, un tel travail a été fait durant la phase de conception entre la première et la seconde version avec un groupe restreint d'utilisateurs (Craheix *et al.*, 2012). La plate-forme PLAGE (Surleau-Chambenoit *et al.*, 2013)¹, avec notamment la mise en place d'un groupe de travail « utilisateurs », offre un lieu d'échanges pérenne entre concepteurs et utilisateurs de méthodes. Mais il ne s'agit pas juste de recueillir des faits et de décrire cette utilisation. Une amélioration passera aussi par la compréhension de la manière dont les utilisateurs perçoivent, comprennent et utilisent l'indicateur ou plutôt ses résultats. Ceci est particulièrement vrai quand on passe aux décideurs et instances politiques. Cette utilisation ne sera pas

forcément directe et peut prendre du temps de familiarisation (Gudmundsson, 2003). En amont, le choix des indicateurs ne se fait pas de manière neutre et peut faire l'objet de sévères débats qui dépassent les faits objectifs comme dans l'exemple des indicateurs de suivi du plan Ecophyto qui ne sont pas indépendants des intérêts de tel ou tel groupe de pression.

La position du chercheur ou des porteurs de la connaissance objective sur les indicateurs aura un rôle différent à jouer selon la nature et le niveau de structuration du débat (Turnhout *et al.*, 2007).

Conclusion

Alors que, dans les années 90, la question de l'existence d'une méthode d'évaluation ou d'un indicateur agri-environnemental pour tel ou tel contexte était primordiale, la multiplication des initiatives dans le domaine amène aujourd'hui à la question du choix pertinent pour l'utilisateur. Des aides existent sous forme de synthèses ou très récemment de sites interactifs comme ceux de PLAGE¹ et GUIDE³. Cependant, l'information sur les méthodes d'évaluation et indicateurs agri-environnementaux reste à élargir à plus de méthodes et à tous les enjeux en allant plus loin que la question de l'azote ou des produits phytopharmaceutiques et de la qualité de l'eau. Elle est aussi à enrichir quant à la qualité prédictive, en ajoutant aussi une analyse qualitative sur le domaine de validité et d'utilisation, comme pour les valeurs de références. Pour le passage à des échelles supérieures, outre le travail méthodologique sur la prise en compte des propriétés émergentes, une question majeure est l'accès aux données. Ceci renforce notre idée de fédérer les compétences et mutualiser les moyens, en termes de données, tout en respectant les intérêts et spécificités des différents partenaires, dans une poursuite d'une plateforme comme PLAGE et/ou du travail du groupe « évaluation multicritère » du GIS GC HP2E. Ceci permettrait de sortir de la situation d'une information très éclatée et diffuse sur les indicateurs et méthodes d'évaluation et sur les aspects méthodologiques qui pourra conduire notamment à un guide méthodologique à la suite d'un premier ouvrage de Girardin *et al.* (2005).

Références bibliographiques

- Assaghir Z., 2010. Analyse formelle et fusion des informations : application à l'estimation et au contrôle d'incertitude des indicateurs agri-environnementaux: Doctorat INPL, Institut National Polytechnique de Lorraine, UMR Nancy-Université - INRA 1121, Nancy, 153 p.
- Bockstaller C., Gaillard G., Baumgartner D., Freiermuth Knuchel R., Reinsch M., Brauner R., Unterseher E., 2006. Méthodes d'évaluation agri-environnementale des exploitations agricoles : Comparaison des méthodes INDIGO, KUL/USL, REPRO et SALCA. Rapport final programme III, projet 4, Colmar, ITADA, 112 p.
- Bockstaller C., Galan M.B., Capitaine M., Colomb B., Mousset J., Viaux P., 2008a. Comment évaluer la durabilité des systèmes en production végétale ? In R. Reau, and T. Doré, (Eds.), Systèmes de culture innovants et durables: quelles méthodes pour les mettre au point et les évaluer: Dijon (France), Educagri, p. 29-51.
- 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. *Agronomy for Sustainable Development* 29, 223-235.
- Bockstaller C., Guichard L., Makowski D., Aveline A., Girardin P., Plantureux S., 2008b. Agri-environmental indicators to assess cropping and farming systems. A review. *Agronomy for Sustainable Development* 28, 139-149.
- Bockstaller C., Lassere-Joulin F., Slezack-Deschaumes, S., Piutti S., Villerd J., Amiaud B., Plantureux S., 2011. Assessing biodiversity in arable farmland by means of indicators: an overview. *Oléagineux Corps gras Lipides* 18, 137-144.

- Bockstaller C., Vertès F., Aarts F., Fiorelli J.L., Peyraud J.L., Rochette P., 2012. Chapitre 8. Méthodes d'évaluation environnementale et choix des indicateurs. In J.L. Peyraud, P. Cellier (coord) *et al.* Les flux d'azote liés aux élevages, réduire les pertes, rétablir les équilibres. Expertise scientifique collective, rapport Inra (France)
http://www.inra.fr/_institut/expertise/expertises_realisees/flux_d_azote_lies_aux_elevages_rapport_d_expertise, p. 335-412.
- Chémery J.-B., Boissier M., 2002. Irri-Mieux : Guide pratique pour l'évaluation d'un projet local, Paris, ANDA, 39 p.
- Colomb B., Bergez J.E., 2013. Construire une image globale des performances des systèmes de cultures par le biais d'une évaluation multicritère. Buts, principes généraux et exemple. *Innovations Agronomiques* 31, 45-60
- Corpen, 2006. Des indicateurs AZOTE pour gérer des actions de maîtrise des pollutions à l'échelle de la parcelle, de l'exploitation et du territoire, Paris, Ministère de l'Ecologie et du Développement Durable, http://www.ecologie.gouv.fr/IMG/pdf/maquette_azote29_09.pdf, 113 p.
- Craheix D., Angevin F., Bergez J. E., Bockstaller C., Colomb B., Guichard L., Reau R., Doré T., 2012. MASC 2.0, un outil d'évaluation multicritère pour estimer la contribution des systèmes de culture au développement durable. *Innovations Agronomiques* 20, 35-48
- Dennis P., Arndorfer M., Balázs K., Bailey D., Boller B., Bunce R. G. H., Centeri C., Corporaal A., Cuming D., Deconchat M., Dramstad W., Elyakime B., Falusi E., Fjellstad W., Fraser M. D., Freyer B., Friedel J. K., Geijzendorffer I., Jongman R., Kainz M., Marcos G.M., Gomiero T., Grausgruber-Gröger S., Herzog F., Hofer G., Jeanneret P., Kelemen E., Kölliker R., Moakes S. R., Nicholas P., Paoletti M. G., Podmaniczky L., Pointereau P., Sarthou J.-P., Siebrecht N., Sommaggio D., Stoyanova S. D., Teufelbauer N., Viaggi D., Vialatte A., Walter T., Widmer F., Wolfrum S., 2009. Conceptual foundations for biodiversity indicator selection for organic and low-input farming systems. Deliverable 2.1, BIOBIO , EU 7th Framework Programme them KBBE-2008-1-2-01, contract no. 227161, <http://www.biobio-indicator.wur.nl/UK/>, 184 p.
- Devillers J., Farret R., Girardin P., Rivière J.-L., Soulas G., 2005. Indicateurs pour évaluer les risques liés à l'utilisation des pesticides, Lavoisier, Londres, Paris, New-York, 278 p, 2-7430-0747-8.
- Galan M.B., Peschard D., Boizard H., 2007. ISO 14 001 at the farm level: Analysis of five methods for evaluating the environmental impact of agricultural practices. *Journal of Environmental Management* 82, 341-352.
- Girardin P., Devillers J., Thybaud E., Soulas G., 2007. Programme "Indicateurs et pesticides" Phase II : Validation et proposition d'amélioration d'indicateurs "pesticides", Ministère de l'Ecologie et du Développement Durable, 71 p.
- Girardin P., Guichard L., Bockstaller C., 2005. Indicateurs et tableaux de bord. Guide pratique pour l'évaluation, Lavoisier, Londres, Paris, New-York, 39 p,
- 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, 184 p
- Gudmundsson H., 2003. The policy use of environmental indicators - learning from evaluation research. *The Journal of Transdisciplinary Environmental Studies* 2, 1-11.
- Halberg N., Van Der Werf H.M.G., Basset-Mens C., Dalgaard R., De Boer I.J.M., 2005. Environmental assessment tools for the evaluation and improvement of European livestock production systems. *Livestock Production Science* 96, 33-50.
- Keichinger O., Benoit P., Boivin A., Bourrain X., Briand O., Chabert A., Domange N., Dubus I. G., Gouy V., Guichard L., Pitrel M., Pleyber E., Roulier S., Zahm F., Bockstaller C., 2013. GUIDE : développement d'un outil d'aide à la sélection d'indicateurs de risques liés à la présence des produits phytopharmaceutiques dans les milieux aquatiques - Mise au point, applications et perspectives. *Innovations Agronomiques* 28, 1-13.
- Leclercq C., Attoumani-Rouceux A., Bockstaller C., Galan M.B., 2011. Usages des méthodes d'évaluation environnementale. Actes du Colloque « Ecologisation des politiques publiques ». <http://www6.paca.inra.fr/ecodeveloppement/Colloque-2011>

- López-Ridaura S., Van Keulen H., Van Ittersum M.K., Leffelaar P.A., 2005. Multi-scale methodological framework to derive criteria and indicators for sustainability evaluation of peasant natural resource management systems. *Environment Development and Sustainability* 7, 51-69.
- Maurizi B., Verrel J.-L., 2002. Des indicateurs pour des actions de maîtrise des pollutions d'origine agricole. *Ingénieries* 30, 3-14.
- Messéan A., Lô-Pelzer E., Bockstaller C., Lamine C., Angevin F., 2010. Outils d'évaluation et d'aide à la conception de stratégies innovantes de protection des grandes cultures. *Innovations Agronomiques* 8, 69-81.
- Niemeijer D., De Groot R.S., 2008. A conceptual framework for selecting environmental indicator sets. *Ecological Indicators* 8, 14-25.
- Oenema O., Kros H., De Vries W., 2003. Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies. *European Journal of Agronomy* 20, 3-16.
- Pierre D., Colmar A., 2009. Réalisation d'une cartographie de l'impact potentiel des produits phytosanitaires sur les eaux superficielles de la région Pays de la Loire – Seconde phase. Annexe 1 Méthodologie SIRIS, Géo-Hyd, http://www.draf.pays-de-la-loire.agriculture.gouv.fr/IMG/pdf/diag_reg_pdl_phyto_annexes_cle4756ba.pdf, 43 p.
- Rosnoblet J., Girardin P., Weinzaepflen E., Bockstaller C., 2006. Analysis of 15 years of agriculture sustainability evaluation methods. In M. Fotyma, and B. Kaminska (Eds.), 9th ESA Congress. Warsaw, Poland, September 4-6, 2006, pp. 707-708.
- Surleau-Chambenoit C., Morin A., Galan M-B., Cariolle M., Leclercq C., Guichard L., Bockstaller C., 2013. PLAGE, un réseau d'acteurs et une plate-forme WEB dédiée à l'évaluation agri-environnementale et de la durabilité des pratiques agricoles, des exploitations agricoles et des territoires. *Innovations Agronomiques* 31, 15-26
- Turnhout E., Hisschemoller M., Eijsackers H., 2007. Ecological indicators: Between the two fires of science and policy. *Ecological Indicators* 7, 215-228.
- Van der Werf H.M.G., Kanyarushoki C., Corson M.S., 2011. L'Analyse de Cycle de Vie : un nouveau regard sur les systèmes de production agricole. *Innovations Agronomiques* 12, 121-133
- Van Der Werf H.M.G., Petit J., 2002. Évaluation de l'impact environnemental de l'agriculture au niveau de la ferme : comparaison et analyse de 12 méthodes basées sur des indicateurs. *Courrier de l'environnement de l'INRA* 46, 121-133.
- Wohlfahrt J., Colin F., Assaghir Z., Bockstaller C., 2010. An indicator inferred from hydrological mechanistic model simulation to assess spatial agricultural management impact on pesticide runoff *Ecological Indicators* 10, 826-839.