



**HAL**  
open science

# Modélisation structurelle : une application à l'évaluation de pratiques écosystémiques au niveau de la parcelle

Dominique Desbois

## ► To cite this version:

Dominique Desbois. Modélisation structurelle : une application à l'évaluation de pratiques écosystémiques au niveau de la parcelle. Cahier des Techniques de l'INRA, 2020, 101, pp.13-32. hal-02635543

**HAL Id: hal-02635543**

**<https://hal.inrae.fr/hal-02635543v1>**

Submitted on 27 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 Attribution - ShareAlike 4.0 International License

## Modélisation structurelle : une application à l'évaluation de pratiques écosystémiques au niveau de la parcelle.

Dominique Desbois<sup>1</sup>



Après plus d'une décennie au service de la Coopération française à l'Université nationale de Côte d'Ivoire comme assistant technique, Dominique Desbois rejoint l'Inra comme responsable informatique de la station de recherche en sciences sociales d'Ivry sur Seine, puis est mis à disposition du service de Statistiques et de Prospective du ministère de l'Agriculture comme chargé de valorisation du Réseau d'informations comptables agricoles. Intégrant l'UMR Économie publique, Dominique Desbois a contribué aux recherches du projet Facepa dans le cadre du 7e programme de Recherche-Développement de l'Union européenne à l'issue duquel il a soutenu une thèse consacrée à l'économétrie des coûts de production. Il a également participé aux projets Trustee, Peerless et Assets de l'ANR, orientés vers l'évaluation des services écosystémiques. Il contribue actuellement à l'évaluation du coût de l'érosion dans le cadre du projet Assess de l'Eranet de coopération euro-méditerranéenne, OTE-Med.

**Résumé.** Cet article propose l'usage de la modélisation structurelle pour l'évaluation de pratiques écosystémiques (exemple contrôle biologique d'agresseurs des cultures) sur la base de données recueillies à l'échelle de la parcelle agricole. Dans la première partie, nous présentons la démarche d'analyse utilisée – modélisation structurelle en moindres carrés partiels, en anglais *Partial Least Squares - Path Modeling (PLS-PM)*. Dans la seconde partie, nous présentons le terrain d'étude et les données considérées dans ce travail. Dans la troisième partie, nous présentons et discutons les résultats issus de la mise en approche *PLS-PM*. Enfin, nous concluons sur la validation de cette approche et les perspectives de son éventuelle extension.

**Mots-clés :** pratiques écosystémiques, modélisation structurelle, *Partial Least Squares*, échelle parcellaire, bio-agresseur, coûts spécifiques, marge brute

**Abstract.** This paper proposes the use of structural modeling for the evaluation of ecosystem-based practices (e.g. biological control of crop pests) on the basis of data collected at the scale of the agricultural plot. In the first part, we present the analytical approach used - structural modeling in partial least squares, i.e. Partial Least Squares - Path Modeling (PLS-PM). In the second part, we present the field of study and the data considered in this work. In the third part, we present and discuss the results from the PLS-PM approach. Finally, we conclude on the validation of this approach and the prospects for its possible extension.

**Keywords:** ecosystem-based practices, structural modeling, *Partial Least Squares*, plot scale, biological pest, specific costs, gross margin

---

<sup>1</sup> INRAE, Economie Publique, 78850 THIVERVAL-GRIGNON, dominique.desbois@inrae.fr

## Introduction

Le contrôle biologique des agresseurs des cultures, ravageurs ou adventices, constitue un exemple de services écosystémiques (SE) – bénéfiques que les écosystèmes fournissent à l'humanité (Millenium Ecosystem Assessment, 2005) – rendu à l'échelle de la parcelle agricole. Les aménités fournies par les infrastructures écosystémiques et les pratiques agro-écologiques en termes de protection du sol, de gestion de la ressource en eau, de préservation de l'habitat des auxiliaires agricoles constituent des services productifs évaluables au niveau de la parcelle et de l'exploitation agricole. Par exemple, le contrôle biologique des agresseurs des cultures, ravageurs ou adventices, par des auxiliaires naturellement présents sur le terrain (comme la coccinelle prédatrice des pucerons) constitue l'un des services productifs mobilisables au niveau de la parcelle pour diminuer la dépendance des systèmes de production agricole aux pesticides. Cependant, les relations entre les modes de gestion des cultures, les niveaux de contrôle des bio-agresseurs par les auxiliaires et les rendements des cultures demeurent encore insuffisamment évalués (Petit *et al.*, 2017). La structure des paysages parcellaires peut également influencer sur le niveau d'abondance des bio-agresseurs, aussi la relation entre paysage parcellaire et rendement reste à inventorier de façon plus exhaustive (Jonsson *et al.*, 2015). L'objectif de ce travail est donc d'analyser via une approche par modélisation structurelle en moindres carrés partiels – en anglais *Partial Least Squares - Path Modeling (PLS-PM)* –, les relations entre : i) paysages parcellaires, ii) ravageurs et adventices des cultures, iii) pratiques agronomiques, et iv) résultats économiques.

La première partie de ce travail présente la spécificité de l'approche *PLS-PM* ; la seconde partie présente le terrain d'étude et décrit les données considérées dans ce travail ; la troisième partie présente les résultats de la mise en œuvre de l'approche *PLS-PM* ; enfin, la dernière partie présente les conclusions sur la validité de cette approche et les perspectives de son application.

## La modélisation par équations structurelles en moindres carrés partiels

La **modélisation par équations structurelles** (*Structural Equation Modeling*) permet principalement d'étudier, via un modèle hypothétique spécifié sous forme d'équations, les liens de causalité (relations) entre plusieurs variables afin de rendre compte du fonctionnement théorique du système étudié (Hoyle, 1995). Dans cette modélisation par équations structurelles, les variables peuvent être soit directement issues d'observations ou de mesures (on parle alors de variables « manifestes »), soit non directement observables (on parle alors de variables « latentes »). Cette approche est très utilisée en sciences sociales (ex. psychologie), où sont étudiés des construits conceptuels non mesurables directement (ex. intelligence, confiance en soi, ...).

## L'approche PLS-PM en moindre carrés partiels

L'approche *PLS-PM* est une variante de la modélisation par équations structurelles qui permet d'analyser un système complexe de relations entre les différentes variables étudiées, sur la base d'un modèle causal *a priori* (*Path Modeling - PM*) décrivant les relations entre variables explicatives ou « exogènes » et les variables expliquées ou « endogènes » (Tenenhaus *et al.*, 2005). La particularité de l'approche *PLS-PM* réside dans le fait que l'estimation des liens du modèle structurel (*path coefficients*) est basée sur le critère d'estimation des moindres carrés partiels (*Partial Least Squares, PLS*), plutôt que celui du maximum de vraisemblance (*Maximum Likelihood - ML*), classiquement utilisé en modélisation par équations structurelles.

Établies par Wold (1981), les propriétés de la **régression des moindres carrés partiels**, *Partial Least Squares (PLS)*, pour l'estimation de systèmes interdépendants ont conduit Lohmöller (1989) à proposer l'approche *PLS* en modélisation par équations structurelles, *PLS-Path Modeling (PLS-PM)*. Ainsi, l'usage de l'approche *PLS-PM* ne nécessite aucune hypothèse sur la distribution des variables (par exemple, normalité de la distribution) et s'avère adapté aux échantillons de faible taille. Les développements théoriques et algorithmiques récents (Tenenhaus *et al.*, 2005) ont ouvert plus largement le champ de son application aux recherches multidisciplinaires où de nombreux groupes de variables sont susceptibles d'interactions conditionnant les phénomènes sociaux ou bien les comportements économiques. En effet, ces recherches multidisciplinaires ne peuvent bien souvent être menées que pour des ensembles de données où les conditions relatives à la normalité de la distribution,

l'indépendance entre observations, ou à la taille de l'échantillon ne sont pas réunies (Chin et Newstead, 1999, p. 314).

La spécification d'un modèle *PLS-PM* comprend les étapes suivantes : i) la spécification d'un modèle hypothétique initial décrivant les relations *a priori* entre les variables latentes et les variables manifestes ; ii) l'estimation des paramètres du modèle via un logiciel statistique approprié ; iii) l'évaluation de la qualité d'ajustement du modèle structurel aux données (*Goodness of Fit - GoF*) ; enfin iv) lorsque cette qualité d'ajustement du modèle est jugée satisfaisante, on procède alors à une dernière étape qui est l'interprétation des résultats.

Dans l'approche *PLS-PM*, le modèle structurel est un ensemble de constructions conceptuelles (ou « variables latentes ») reliées par des relations causales hypothétiques (le « modèle interne ») qui peuvent être estimées au moyen de « variables manifestes » mesurées ou observées reflétant ou, respectivement, formant les variables latentes (modèle externe). La figure 1 présente un exemple de modélisation structurelle spécifiant le modèle interne et le modèle externe.

### Spécification du modèle interne

Les relations structurelles entre variables latentes « endogènes » (expliquées) et « exogènes » (explicatives) constituent le modèle interne et sont formalisées par les équations linéaires suivantes :

$$LV_l^{endo} = \beta_l^0 + \sum_{k=1}^K \beta_l^k LV_k^{exo} + \varepsilon_l$$

où  $\beta_l^k$ , appelé « lien structurel » (*path coefficient*), représente le signe et la force de la relation déterministe entre la variable latente endogène  $LV_l^{endo}$  et les variables latentes exogènes  $LV_k^{exo}$ . La partie non expliquée par le modèle déterministe de  $LV_l$  est reléguée au sein du résidu  $\varepsilon_l$ . Les liens structurels sont estimés par une procédure de projection géométrique (moindres carrés) dont la seule hypothèse est l'indépendance entre la partie déterministe et la partie résiduelle, imposant que la covariance entre chaque variable latente exogène et le résidu soit nulle :  $cov(LV_k^{exo}, \varepsilon_l) = 0$ .

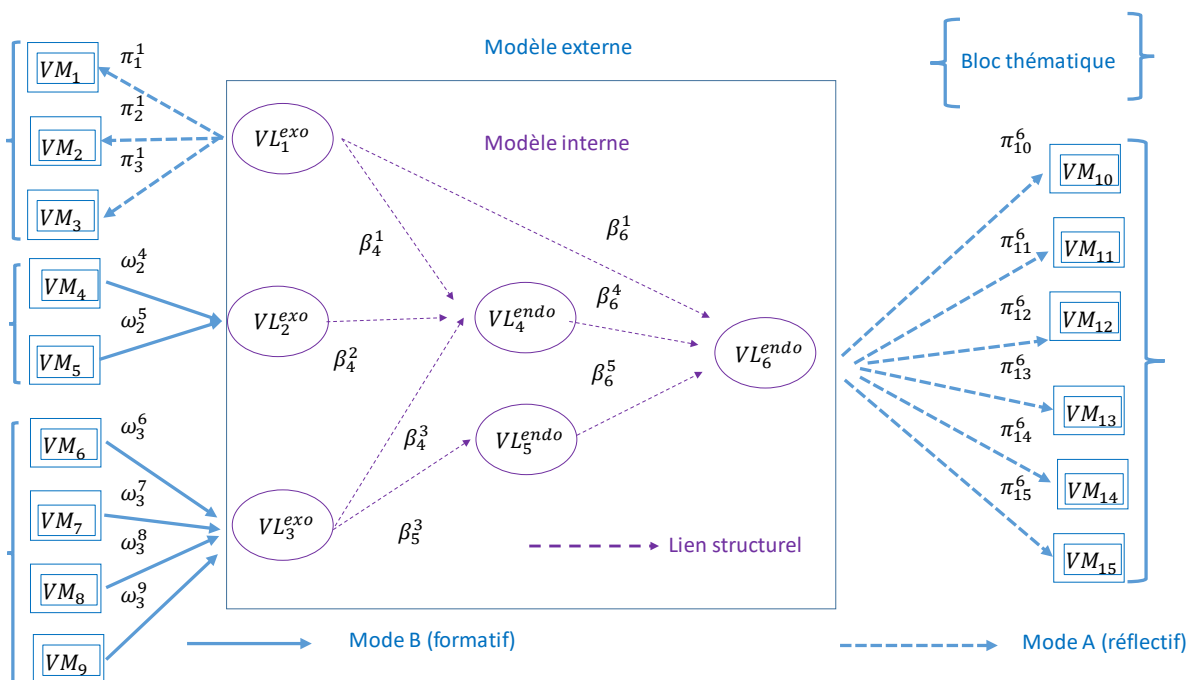


Figure 1. Exemple de modèle structurel décrivant les relations entre variables latentes et variables manifestes

**Lecture** : les coefficients **beta** représentent les « liens » du modèle interne ; les variables manifestes  $VM_h$  sont associées aux variables latentes exogènes ( $VL_k^{exo}$ ) ou endogènes ( $VL_k^{endo}$ ) selon un mode qui peut être soit « **réflectif** » (mode **A**) : les variables manifestes « reflètent » les variables latentes, soit « **formatif** » (mode **B**) : les variables manifestes « forment » les variables latentes).

### Spécification du modèle externe

Les variables latentes sont définies par les « variables manifestes » (VM), issues de mesures et/ou d'observations qu'elles soient directes ou indirectes, via deux modes (cf. figure 1) : i) le mode « réflectif » où les variables latentes se reflètent à travers leurs effets ou conséquences sur les indicateurs observés ; le mode « formatif » où les variables latentes sont censées être formées ou constituées par les variables mesurées.

Dans le **mode réflectif** (A), la variable manifeste  $X_k$  reflète la variable latente  $VL_k$ , de moyenne  $m$  et d'écart-type 1, selon le schéma projectif suivant (régression des moindres carrés) :

$$X_h = \pi_h^0 + \pi_h^k VL_k + \varepsilon_h$$

où le coefficient  $\pi_h^k$  est le « **poids externe** » (**outer weight**) de la variable latente influant la variable manifeste. Le résidu  $\varepsilon_h$  est de moyenne nulle et indépendant de la variable latente ( $cov[VL_k, \varepsilon_h] = 0$ ).

Dans le **mode formatif** (B), les variables mesurées « forment » les variables latentes, selon l'équation suivante :

$$LV_k = \sum_{h=1}^H \omega_k^h X_h + \delta_k$$

où le coefficient  $\omega_k^h$  est un « **facteur structurel** » (**structural loading**) contributif à la variable latente. Le résidu  $\delta_k$  est de moyenne nulle et indépendant de chacune des variables manifestes ( $cov[X_h, \delta_k] = 0$ ).

Le mode le plus couramment utilisé est le **mode réflectif** (mode A).

Dans l'estimation des paramètres du modèle, l'approche PLS-PM vise à maximiser globalement la variance expliquée des variables endogènes.

### Statistiques de validation du modèle externe

L'unidimensionnalité du bloc de variables manifestes correspondant à chaque variable latente est une hypothèse structurelle du modèle externe qu'il convient de pouvoir valider à l'aide des différents critères présentés ci-après.

#### i) Écart entre les deux premières valeurs propres du bloc de données

Premier critère d'unidimensionnalité, l'analyse en composantes principales du bloc de données correspondant à chacune des variables latentes (cf. tableau 1) fournit un premier critère d'unidimensionnalité adapté de la règle de Kaiser : si la première valeur propre de la matrice des corrélations est supérieure à 1 et la seconde valeur propre est beaucoup plus petite, cela signifie que la très grande majorité des variables manifestes est positivement corrélée à la première composante principale.

#### ii) Alpha de Cronbach

Le second critère d'unidimensionnalité est l'**alpha de Cronbach**, rapport de la somme des covariances sur la variance de la somme des  $H$  variables manifestes du bloc de données correspondant à une variable latente,

$$\text{soit : } \alpha = \frac{\sum_{h=1}^H cov(X_h, X_{h'})}{var(\sum_{h=1}^H X_h)} \times \frac{H}{H-1}$$

L' $\alpha$  de Cronbach est très utilisé dans les analyses de fiabilité avec la règle suivante : si ce rapport est supérieur à 0,7, alors le bloc peut être considéré comme unidimensionnel.

## Le Cahier des Techniques de l'Inra 2020 (101)

### iii) Rho de Dillon-Goldstein

Le dernier critère utilisé est le *rho de Dillon-Goldstein*, rapport de la variance de la variable latente à la variance de son bloc de variables manifestes,

estimable par  $\hat{\rho} = \frac{[\sum_{h=1}^H \text{corr}(X_h, t_1)]^2}{[\sum_{h=1}^H \text{orr}(X_h, t_1)]^2 + \sum_{h=1}^H (1 - [\text{orr}(X_h, t_1)]^2)}$  où  $t_1$  est la première composante principale du bloc thématique de variables manifestes.

Si l'estimation du  $\rho$  de Dillon-Goldstein est supérieur à 0,7, alors le bloc est considéré comme unidimensionnel.

Le  $\rho$  de Dillon-Goldstein est considéré comme un critère meilleur que l' $\alpha$  de Cronbach par Chin (1998) car il est fondé sur les facteurs structurels du modèle interne, plutôt que sur les corrélations entre les variables manifestes du modèle externe faisant implicitement l'hypothèse de variables manifestes *a priori* équivalentes entre elles pour définir une variable latente (hypothèse de la  $\tau$ -équivalence).

### iv) Communalité

La « **communalité** »  $com_k$  du  $k^e$  bloc thématique indique dans quelle mesure la variabilité des variables manifestes du  $k^e$  bloc est restituée par les scores de la  $k^e$  variable latente. La « communalité » du bloc thématique  $k$  est égale à la somme pondérée des carrés des corrélations entre les variables manifestes et la variable latente centrée réduite  $Y_k$ , soit :

$$com_k = \frac{1}{H_k} \sum_{h=1}^{H_k} cor^2(X_h, Y_k).$$

## Validation globale de la modélisation structurelle

### L'indice moyen de redondance

Afin de relier la performance prédictive du modèle externe de mesure à la cohérence des composantes du modèle interne, l'**indice de redondance** calculé pour chaque bloc thématique endogène mesure la part de variabilité des variables manifestes reliées aux variables latentes  $Y_h$  expliquant la  $k^e$  variable latente endogène, soit :

$$Red_k = com_k \times R^2(Y_k^{endo}, Y_{h:Y_h \rightarrow Y_k^{endo}})$$

Il s'interprète comme un indicateur de la capacité à prédire les valeurs observées de la  $k^e$  variable latente endogène.

L'indice moyen de redondance,  $\overline{Red}$ , calculé sur l'ensemble des variables endogènes, soit :

$$\overline{Red} = \frac{1}{K^{endo}} \sum_{k=1}^{K^{endo}} Red_k$$

fournit ainsi un indice global de la capacité de prédiction des valeurs observées des variables latentes endogènes du modèle.

### La qualité d'ajustement

Proposé par Amato *et al.* (2004), la **qualité d'ajustement** (*Goodness of Fit*, abrégé en **GoF**) du modèle est définie par la racine carrée du produit de la « communalité » moyenne par le  $R^2$  moyen, soit :

$$GoF = \sqrt{\overline{com} \times R^2} = \sqrt{\frac{\sum_{k=1}^K \sum_{h=1}^{H_k} corr^2(X_k^h, Y_k)}{\sum_{k=1}^K H_k} \times \frac{\sum_{k=1}^{K^{endo}} R^2(Y_k^{endo}, [Y_j \text{ explicatives de } Y_k^*])}{K^*}}$$

où  $\overline{com}$ , la « communalité moyenne » est la moyenne pondérée des communalités de chacun des blocs thématiques, soit :  $\overline{com} = \frac{1}{\sum_{k: H_k > 1} H_k} \sum_{k: H_k > 1} H_k com_k$

Puisqu'au sein de chaque bloc les communalités thématiques sont des moyennes de carrés de coefficients de corrélation, la communalité moyenne est la moyenne de l'ensemble des carrés de coefficient de corrélation entre les variables latentes et leurs variables manifestes.

### L'amorçage

Puisque l'approche *PLS-PM* ne fait pas appel à des hypothèses distributionnelles, le recours aux procédures de validation basées sur le *bootstrap* ou « amorçage » (Efron et Tibshirani, 1993) devient nécessaire dans une approche non plus seulement exploratoire mais confirmatoire. Des estimations amorcées peuvent être calculées pour les poids externes, les facteurs, les liens structurels, les indicateurs de communalité, de redondance, et la qualité d'ajustement global du modèle.

Le principe de la procédure d'amorçage est de tirer aléatoirement  $B$  nouveaux échantillons ( $B < 100$  habituellement) de  $N$  observations (les « amorces ») au sein de l'échantillon initial des observations pour avoir une estimation de la fonction quantile  $\varphi$ , inverse de la fonction de répartition (distribution des probabilités cumulées). Par exemple, les valeurs « amorcées » des liens structurels sont estimées à partir des  $B$  amorces sur la base d'une procédure de Monte-Carlo, fournissant les intervalles de confiance empiriques estimés de niveau  $(1 - \tau)$  de la fonction quantile des communalités amorcées  $\hat{\varphi}_{lien}^B$ , réciproque de la fonction de répartition, soit :  $[\hat{\varphi}_{lien}^B(\tau/2); \hat{\varphi}_{lien}^B(1 - \tau/2)]$ .

## Matériel et méthode

### Contexte agro-écologique de l'étude

Dans cette étude, nous appliquons l'approche *PLS-PM* au contexte agro-écologique défini par les parcelles de terrains expérimentaux ou agricoles suivies dans le cadre de recherches portant sur quatre zones d'études (ZE) (cf. figure 3). Ces recherches sont fédérées par le projet multidisciplinaire « *Predictive Ecological Engineering for Landscape Ecosystem Services and Sustainability* » (*Peerless*), et financées par l'Agence Nationale de la Recherche (ANR).



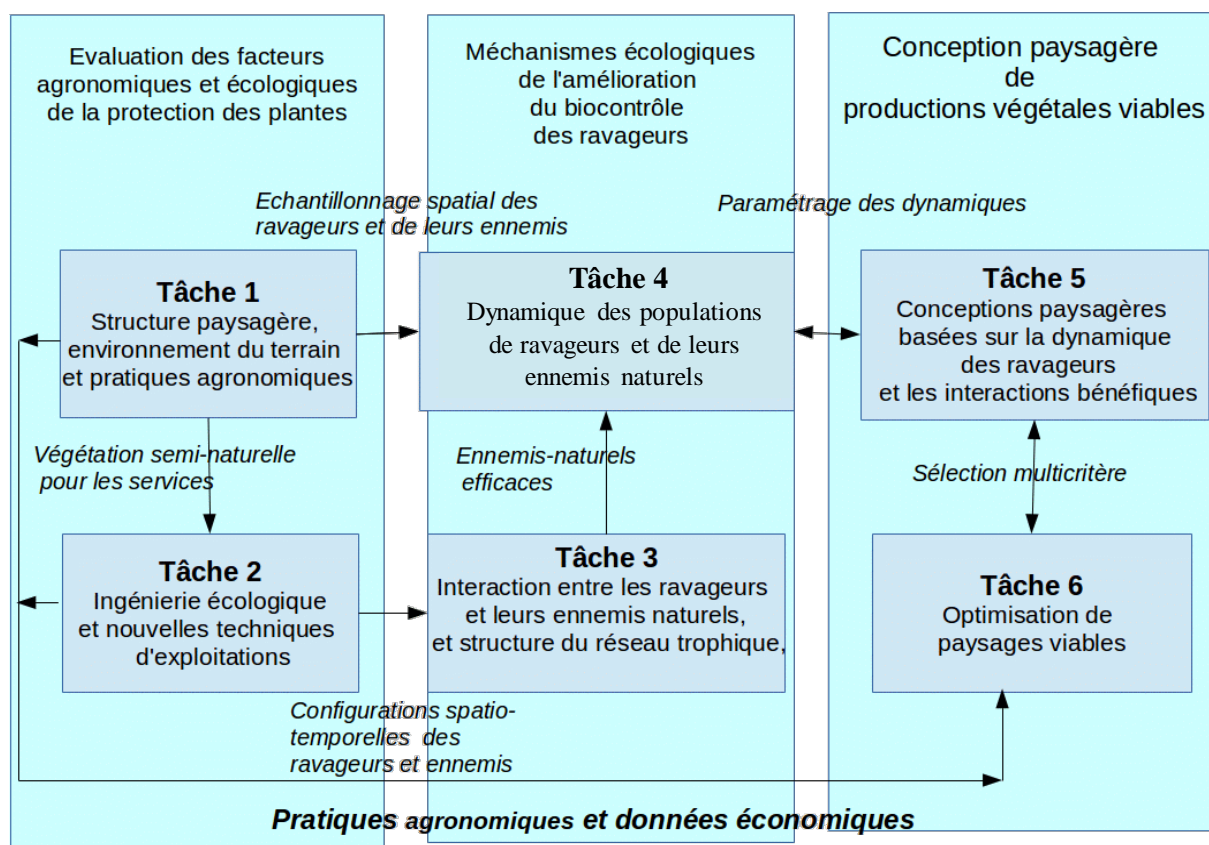


Figure 2. Infrastructures et pratiques agro-écologique du projet multidisciplinaire Peerless.

Source : d'après Franck *et al.* (2017).

Le projet *Peerless* vise à identifier des stratégies de gestion alternatives améliorant les pratiques de lutte phytosanitaire en s'appuyant sur la biodiversité fonctionnelle en arboriculture et en grandes cultures pour optimiser les systèmes de production agricoles, aux échelles locales et paysagères, dans une double perspective de durabilité et de viabilité économique de ces productions (Franck *et al.*, 2017). *Peerless* a été structuré autour des trois objectifs suivants : i) l'évaluation des facteurs agronomiques et écologiques de la protection des plantes ; ii) l'identification des mécanismes écologiques de l'amélioration du bio-contrôle des ravageurs ; iii) la conception paysagère de productions végétales viables. Le projet *Peerless* agrège quatre zones d'étude (cf. figure 3) en grandes cultures (Anjou-Saumurois, Armorique, Côte d'Or et Île de France), et une zone d'étude en arboriculture (Basse Vallée de la Durance). Ces zones d'études (ZE) représentent un ensemble de 158 parcelles enquêtées sur les années 2014 et 2015.



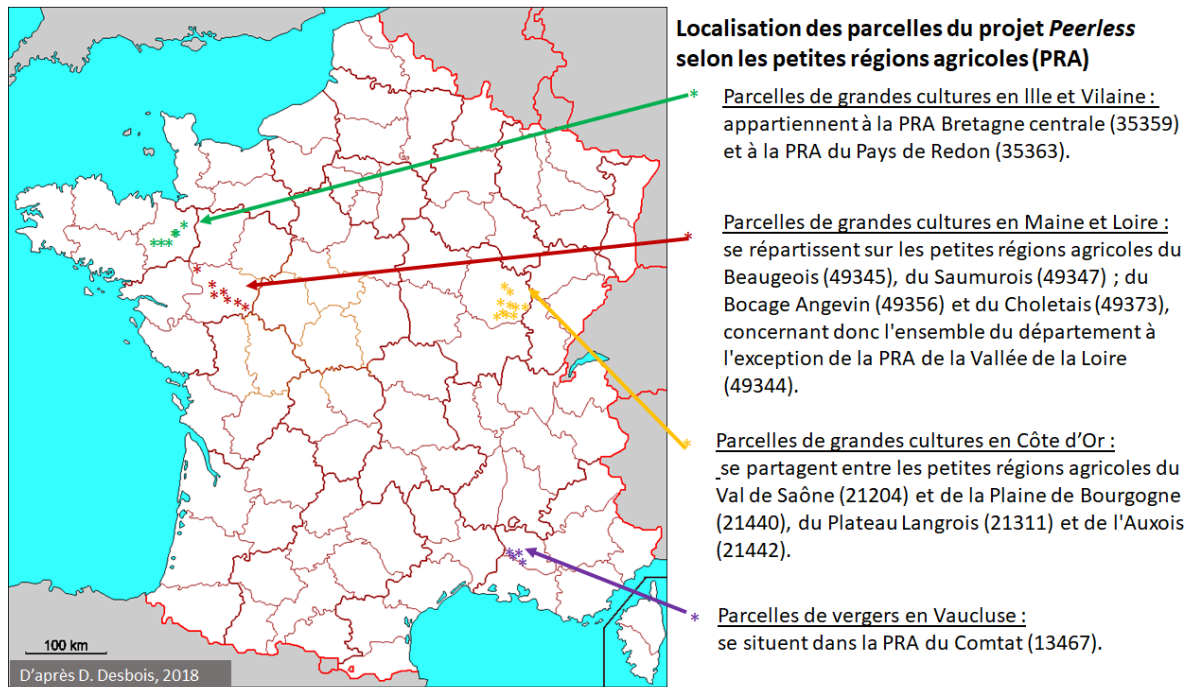


Figure 3. Localisation des zones d'étude du projet multidisciplinaire *Peerless*.

Source : d'après Desbois (2018).

### Collecte de données

Dans ce travail, les données du projet *Peerless* considérées à l'échelle des différentes ZE sont : i) des mesures agro-écologiques de terrain ; ii) des observations de terrain sur les pratiques agronomiques ; iii) des estimations économiques basées sur les pratiques et rendements parcellaires et les références de prix généralement régionales, sinon nationales.

- Les **données agro-écologiques** ont été collectées dans le cadre de la tâche T1 (*Structure paysagère, environnement du terrain et pratiques agronomiques*) du projet. Pour chacune des zones d'étude, la caractérisation du paysage a été effectuée au sein d'un cercle de surface unitaire de 1km<sup>2</sup> selon deux modalités. Première caractérisation du paysage (*Paysage1*) : les contextes paysagers des parcelles sont caractérisés au moyen des proportions (pourcentages) en cultures d'hiver, cultures de printemps, cultures d'été, cultures pérennes, fabacées, friche, maraîchage, oléagineux, et des proportions en prairies et en zones boisées. Seconde caractérisation du paysage (*Paysage2*), deux indices de Shannon ont été calculés en prenant le plus bas niveau de correspondance des nomenclatures entre les différents sites : i) le premier calculé sur les grandes cultures ; ii) le second calculé sur l'ensemble des milieux étudiés (cultures, prairies et zones boisées).
- Les **pratiques agronomiques** sont documentées par les indicateurs des itinéraires techniques culturaux, à savoir : le nombre de passages en labour, et en travail du sol (profond et superficiel) ; le nombre de fertilisations en général et azotées (N) en particulier, qu'elles soient minérales ou organiques ; les dates de semis et de récolte ; ainsi que l'ensemble des indicateurs de fréquence de traitement (IFT), qu'ils soient en mode d'exploitation conventionnel ou biologique.
- Les **données économiques** ont été produites dans le cadre de la tâche T6 (*Optimisation de paysages viables*) du projet et fusionnées avec les données agro-écologiques (Desbois, 2018) en mobilisant à la fois les références de prix issues de la statistique publique, celles produites par les instituts techniques (Arvalis, Centre technique interprofessionnel des Fruits et Légumes) et les établissements publics (FranceAgriMer, Chambres d'Agriculture) pour le calcul des produits bruts, des coûts spécifiques et des marges.

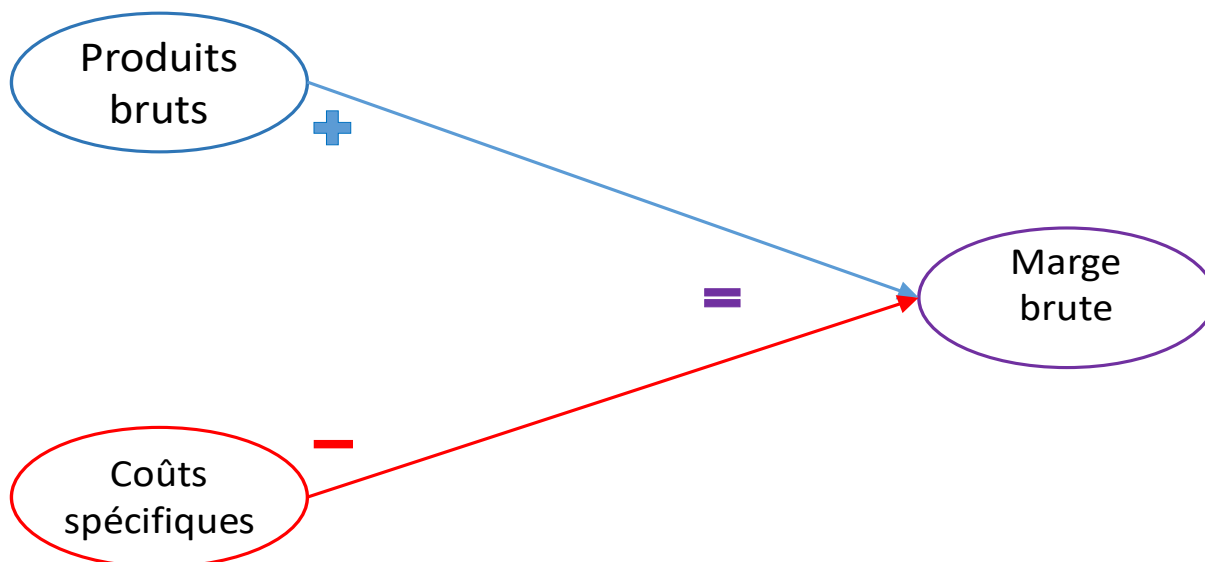


Figure 4. Modèle interne simplifié de la marge brute de la parcelle cultivée

Lecture : le symbole (+) dénote l'ajout des produits bruts et celui (-) de la déduction des coûts spécifiques.

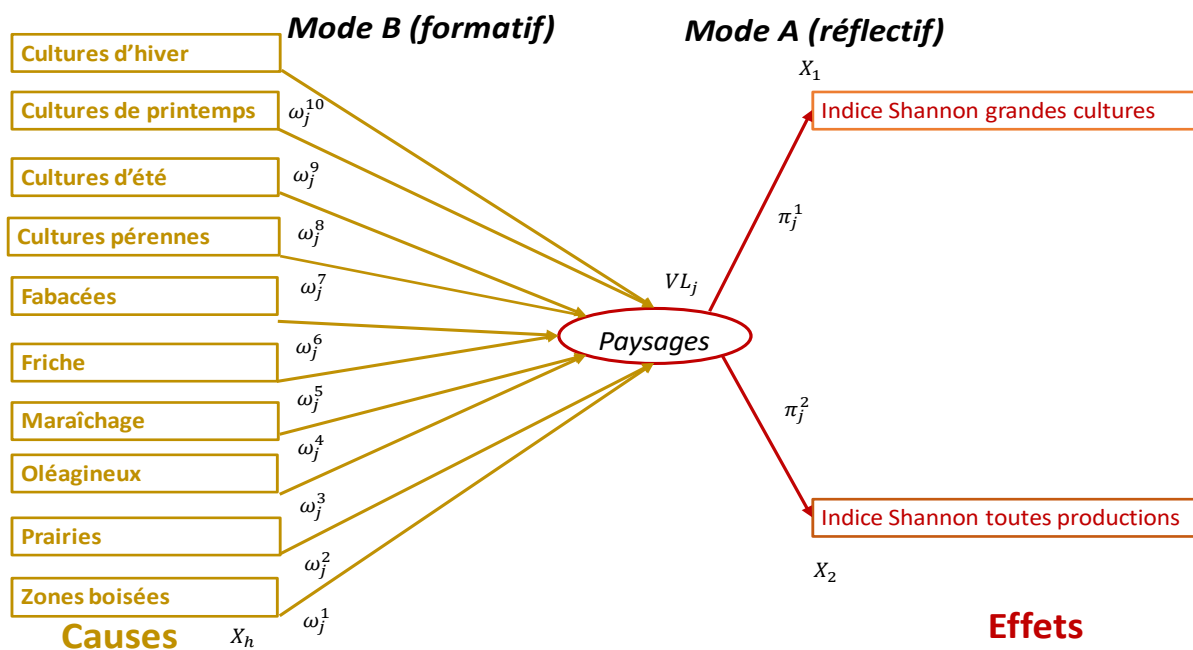


Figure 5. Les modes formatif et réflectif des relations entre variables manifestes et variables latentes.

Lecture : en marron, les variables manifestes en mode formatif (causes) et leurs facteurs structurels ( $\rightarrow$ ); en bistre, les variables manifestes en mode réflectif (effets) et leurs poids externes ( $\leftarrow$ ).

## Le modèle structurel et son estimation

Construite à partir des concepts de la comptabilité des entreprises, la relation comptable simplifiée  $MargeBrute = ProduitBrut - CoûtsSpécifiques$  définissant le concept de marge brute comme somme algébrique du produit brut et des coûts spécifiques, offre un exemple simple d'application non seulement aux exploitations agricoles mais également aux parcelles cultivées de la notion de modèle structurel interne.

Une fois spécifié, ce modèle interne permet de structurer l'ensemble des variables mesurées ou observées (les « variables manifestes ») en plusieurs blocs correspondant aux construits conceptuels, chaque bloc de variables manifestes représentant une variable latente. La composition précise en variables manifestes des blocs thématiques correspondant aux différentes variables latentes (*Paysage1*, *Paysage2*, *Ravageur*, *Adventic*, *Costspec*, *Prodbrut*, *Subsidys*, *Margeuha*) figure en annexe au tableau 5.

Dans un contexte d'application agro-écologique, le concept de *Paysages* de la parcelle agricole peut être constitué en mode formatif (B) par les profils d'occupation des sols (les différentes cultures, les fabacées, la friche, le maraîchage, les oléagineux, les prairies et zones boisées) comme il peut se refléter (mode A) dans les différents indices de diversité qui peuvent être construits à partir de sa description (indices de Shannon « grandes cultures » ou « toutes productions »).

Appliqué aux contextes écosystémiques résultant de la conjonction d'infrastructures et de pratiques, la spécification d'un modèle structurel d'interaction agro-écologique et d'impact économique (figure 5) permet d'analyser les interrelations entre infrastructures écologiques (profils de paysage au niveau de la parcelle), pratiques agronomiques (labour, cultures intermédiaires), et résultats induits (augmentation du produit, réduction des coûts).

Dans ce travail, le modèle structurel initial reprend celui proposé par Mezerette (2016), en le complétant par les variables économiques de coûts, de subventions, de produit, et de marge (cf. figure 1) et en étendant son application aux sites arboricoles de *Peerless*. Ce modèle structurel initial explicite les relations *a priori* entre les variables agro-écologiques décrivant le paysage (*Paysage1*, *Paysage2*), les pratiques agronomiques (*Practice*) ainsi que la présence de bio-agresseurs, ravageurs (*Ravageur*) et adventices (*Adventic*), avec les variables économiques de coûts spécifiques (*Costspec*), de production brute (*Prodbrut*), de soutiens à la production (*Subsidys*), et enfin de marge brute à l'hectare (*Margeuha*) au niveau de la parcelle.

Dans le schéma de la figure 5, les infrastructures écologiques (*Paysage1*, *Paysage2*) influent à la fois sur les ravageurs (*Ravageur*), les adventices (*Adventic*) et les pratiques agronomiques (*Practice*), dans un sens *a priori* indéterminé ( $\rightarrow$ ). Les pratiques agronomiques impactent *a priori* les produits bruts (*Prodbrut*) et les coûts spécifiques (*Costspec*) soit directement en les renforçant ( $\rightarrow$ ), soit indirectement en affaiblissant ( $\rightarrow$ ) les ravageurs et les adventices. Les ravageurs et les adventices sont supposés avoir un impact dépressif ( $\rightarrow$ ) sur les produits bruts et inflationniste ( $\rightarrow$ ) sur les coûts spécifiques. Le solde de ces influences sur les produits bruts et les coûts spécifiques détermine, avec les subventions (*Subsidys*), l'impact *in fine* des pratiques agro-écologiques sur la marge brute unitaire à l'hectare (*Margeuha*). Dans ce schéma, les construits conceptuels *Paysage1*, *Paysage2*, et *Subsidys* apparaissent comme des VL *a priori* exogènes (*i.e.* non déterminées par d'autres phénomènes) tandis que les concepts *Practice*, *Ravageur*, *Adventic*, *Costspec*, *Prodbrut*, et *Margeuha* sont des VL *a priori* endogènes (*i.e.* influencées par d'autres phénomènes).

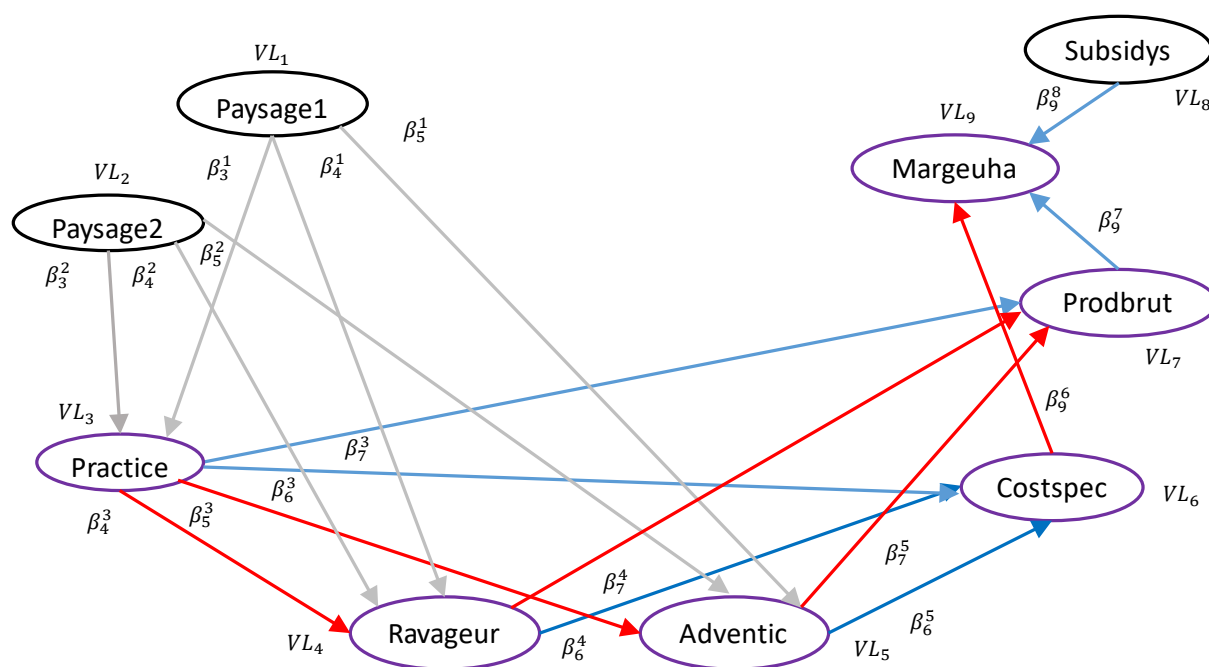


Figure 6. Le modèle structurel d'interaction agro-écologique et d'impact économique au niveau de la parcelle.

Lecture du graphique : en noir, les VL exogènes ; en violet, les VL endogènes ; en bleu, les liens structurels de renforcement ; en rouge, les liens structurels d'affaiblissement ; en gris, les liens structurels a priori indéterminés.

## Résultats et discussion

Le modèle structurel initial est estimé à partir d'un ensemble de 158 parcelles enquêtées au cours des années 2014 et 2015. La mise en œuvre de l'approche PLS-PM a été réalisée via le logiciel statistique **R version 3.6.2** (<https://www.r-project.org/>), et principalement via le package « *pls-pm* », version 0.4.9 (Sanchez, Trinchera et Russolillo, 2017). L'estimation des liens structurels (*path coefficients*) du modèle a été conduite en utilisant le mode réflectif (A) pour l'ensemble des variables latentes.

La figure 6 permet d'avoir une vue d'ensemble du modèle externe. Par exemple, pour la variable latente *Ravageurs*, les variables manifestes reflétant cette variable latente au sein du bloc thématique correspondant sont : *s.avenae*, *m.dirhodum*, *Total.pucerons*, *Larves.de.Lema*, *Pctdagats..Lema*. La valeur des facteurs structurels est positive pour les dénombrements de ravageurs (*Total.pucerons*, *Larves.de.Lema*) et l'indicateur de leurs dégâts (*Pctdagats..Lema*). Cependant, cette valeur est négative pour les dénombrements du sitobion (*s.avenae*), puceron spécifique des épis de céréales, à l'inverse des autres dénombrements de ravageurs que sont le puceron des céréales et du rosier (*m.dirhodum*) et les criocères des céréales (*lema*).

## Vérification de l'unidimensionnalité de chaque bloc

Les statistiques d'unidimensionnalité (alpha de Cronbach et rho de Dillon-Goldstein) réalisées à partir du modèle structurel initial présenté (cf. figure 6) montrent des valeurs trop faibles pour les blocs thématiques *Subsidys* (Subventions), *Paysage2*, et *Costspec* (cf. tableau 1). En effet, pour ces blocs, on observe un certain nombre de corrélations (*loadings*) négatives, signalées par des flèches rouges au sein du schéma initial de la modélisation structurelle (cf. figure 7).

Tableau 1. Statistiques d'unidimensionnalité des blocs thématiques du modèle initial.

Blocs	Mode	Variables manifestes	Alpha de Cronbach	Rho de Dillon-Goldstein	Valeur propre 1	Valeur propre 2
<i>Paysage1</i>	A	10	0,000	0,215	2,100	1,789
<i>Paysage2</i>	A	2	0,829	0,921	1,710	0,292
<i>Practice</i>	A	19	0,720	0,741	9,110	3,331
<i>Ravageur</i>	A	5	0,717	0,838	3,030	0,962
<i>Adventic</i>	A	6	0,871	0,904	3,680	1,090
<i>Costspec</i>	A	10	0,541	0,611	3,020	2,496
<i>Prodrbrut</i>	A	4	0,592	0,790	2,300	1,588
<i>Subsidys</i>	A	4	0,664	0,804	2,080	1,309
<i>Margeuha</i>	A	3	0,986	0,991	2,920	0,066

Lecture : les valeurs en *bleu* indiquent que le critère choisi d'unidimensionnalité est satisfait.

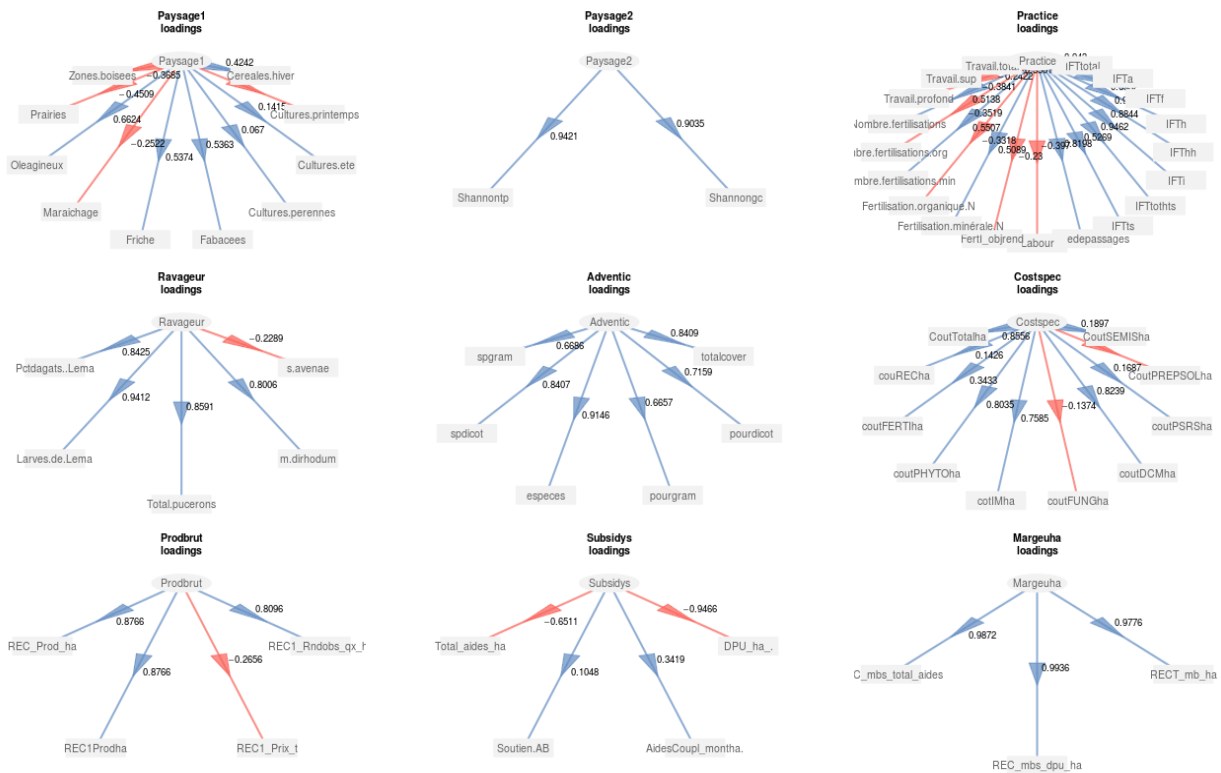


Figure 7. Modèle externe initial de la modélisation structurelle

Lecture : les flèches en *bleu* (→) indiquent un facteur structurel de la variable manifeste reflétant positivement la variable latente, les flèches *rouges* (→) négativement, comme l'indique la valeur adjacente du facteur structurel.

Les blocs *Paysage2*, *Practice*, *Ravageur*, *Adventic* et *Margeuha* sont considérés comme unidimensionnels car ils satisfont tous les critères. Le bloc *Paysage1* et les blocs *Costspec*, *Prodrbrut*, et *Subsidys* présentent des valeurs inférieures à 0,7 pour l'alpha de Cronbach, cependant les blocs *Prodrbrut* et *Subsidys* peuvent être considérés comme unidimensionnels car ils satisfont au critère du rho de Dillon-Goldstein.

## Évaluation de la qualité de la modélisation et ajustement du modèle externe

La qualité d'ajustement (GoF) du modèle initial est de 0,4285. Cet indicateur peut être amélioré par ajustement du modèle externe : en effet, certaines liaisons corrélées négativement peuvent par un recodage approprié être transformées en liaisons positives. Ainsi, dans le bloc *Subsidys*, le recodage de la variable manifeste (VM) *AidesCoupl\_montha* en la variable  $PctAideDecoupl = Total\_aides\_ha - AidesCoupl\_montha$  permet d'obtenir une corrélation positive (cf. figure 8). De même, pour le bloc thématique *Paysage1*, la VM *Maraichage* est recodée en la variable  $TotCultNonMaraicher = 100 - Maraichage$ . Pour le bloc *Pratique*, la VM *IFTf* est transformée en  $TotIFTNonFung = IFTtotal - IFTf$ . Pour le bloc *Costspec*, la VM *coutFUNGha* est transformée en la variable  $DifTotcoutFung = CoutTotalha - coutFUNGha$ .

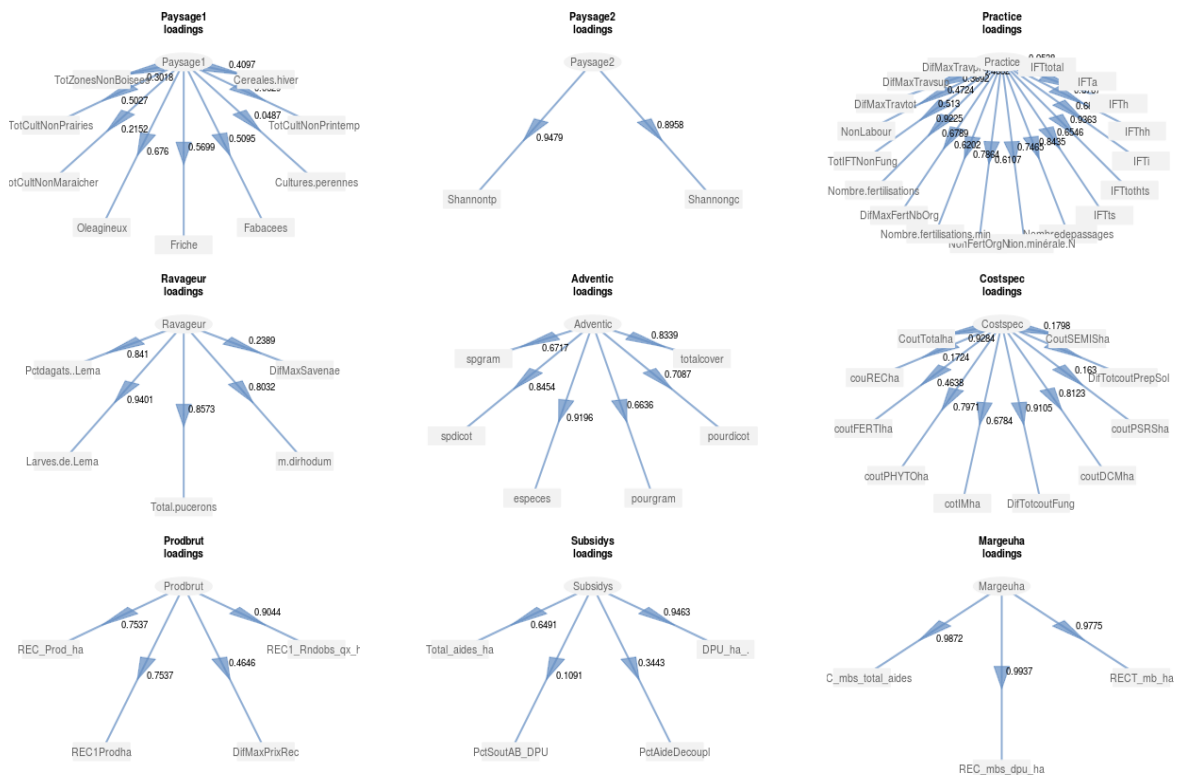


Figure 8. Modèle externe révisé après ajustement

**Lecture :** les flèches en **bleu** (→) indiquent un facteur structurel de la variable manifeste reflétant positivement la variable latente.

Après ce recodage, les statistiques d'unidimensionnalité se sont améliorées (cf. tableau 2) : elles sont toutes supérieures à 0,7 (seuil d'acceptation de l'hypothèse d'unidimensionnalité) sauf pour les blocs *Paysage1* et *Prodrut*, le bloc *Subsidys* avec des valeurs inférieures à 0,5 pour l'alpha de Cronbach, cependant proches du seuil d'acceptation pour *Paysage2* avec le rho de Dillon-Goldstein (>0,6). La valeur du GoF, l'indicateur global de la qualité d'ajustement du modèle amendé s'améliore de façon assez marginale, passant à 0,5078 (cf. tableau 3).

## Modèle structurel final après révision

La prise en compte des corrélations entre blocs aboutit au modèle structurel final suivant (cf. figure 4) où les relations entre variables latentes apparaissent du signe attendu sauf pour la variable latente « Adventices » issu du bloc *Adventic* présentant une relation négative quoique faible avec la variable latente « Coûts spécifiques » issue du bloc *Costspec*.

Remarquons que la relation entre la variable latente *Practice* représentant les pratiques agronomiques et la variable latente *Ravageur* représentant la présence de ravageurs est positive alors qu'elle est négative avec la variable latente *Adventic* représentant la présence d'adventices. Il pourrait donc y avoir intérêt à distinguer les pratiques spécifiques du contrôle des ravageurs de celles spécifiques du contrôle des adventices.

Tableau 2. Statistiques d'unidimensionnalité des blocs thématiques du modèle amendé.

Bloc thématique	Mode	Variables manifestes	Alpha de Cronbach	Rho de Dillon-Goldstein	Valeur propre 1	Valeur propre 2
<i>Paysage1</i>	A	9	0,398	0,073	2,040	1,673
<i>Paysage2</i>	A	2	0,829	0,921	1,710	0,292
<i>Practice</i>	A	18	0,943	0,950	9,450	3,201
<i>Ravageur</i>	A	5	0,807	0,874	3,030	0,962
<i>Adventic</i>	A	6	0,871	0,904	3,680	1,090
<i>Costspec</i>	A	10	0,834	0,879	4,750	2,301
<i>Prodbrut</i>	A	4	0,690	0,805	2,300	1,588
<i>Subsidys</i>	A	4	0,243	0,125	2,090	1,313
<i>Margeuha</i>	A	3	0,986	0,991	2,920	0,066

Lecture : les valeurs en bleu indiquent que le critère choisi d'unidimensionnalité est satisfait selon les règles empiriques définies pour chaque indicateur.

Tableau 3. Communalités, R<sup>2</sup> et redondances.

Bloc thématique	Variables manifestes	Communalité	Type	R <sup>2</sup>	Redondance moyenne
<i>Paysage1</i>	9	0,1780	Exogène	0,0000	0,0000
<i>Paysage2</i>	2	0,8500	Exogène	0,0000	0,0000
<i>Practice</i>	18	0,5170	Endogène	0,0921	0,0477
<i>Ravageur</i>	5	0,6060	Endogène	0,1254	0,0759
<i>Adventic</i>	6	0,6080	Endogène	0,2806	0,1707
<i>Costspec</i>	10	0,4620	Endogène	0,6764	0,3127
<i>Prodbrut</i>	4	0,5420	Endogène	0,6906	0,3746
<i>Subsidys</i>	4	0,3620	Exogène	0,0000	0,0000
<i>Margeuha</i>	3	0,9720	Endogène	0,6391	0,6215
<i>Moyenne pondérée</i>		0,4990			
<i>Moyenne des bloc endogènes</i>		0,6178		0,4174	0,2579
<i>Qualité d'ajustement (GoF)</i>				0,5078	

Lecture : la qualité de l'ajustement (Goodness of Fit - GoF) est calculée à partir des blocs endogènes.

Les effets directs et indirects sont présentés en figure 10 sous forme de diagramme en bâtons, le cas échéant empilés lorsqu'ils se cumulent. Concernant les effets, directs ou indirects, nous distinguons ceux importants (supérieurs à 0,4), de ceux qui sont modérés (supérieurs à 0,2 et inférieurs à 0,4) ou qui sont faibles (inférieurs à 0,2 et supérieurs à 0,1), voire très faibles (inférieurs à 0,1).

Parmi les effets directs importants, figurent : i) les effets négatifs des pratiques agronomiques sur les adventices ; ii) les effets négatifs des ravageurs et des adventices sur le produit ; iii) et les effets positifs du produit sur la



## Le Cahier des Techniques de l'Inra 2020 (101)

marge brute. Les effets indirects les plus importants (supérieurs à 0,4) sont ceux négatifs des ravageurs sur la marge brute.

Tableau 4. Estimation amorcée des liens structurels du modèle interne.

Lien	Estimation	Estimation	Erreur-type	Q(0,025)	Q(0,975)
		eustachisée			
<i>Paysage1 -&gt; Practice</i>	0,305	0,249	0,323	-0,521	0,639
<i>Paysage1 -&gt; Ravageur</i>	0,115	0,203	0,251	-0,380	0,552
<i>Paysage1 -&gt; Adventic</i>	0,398	0,431	0,215	-0,227	0,692
<i>Paysage2 -&gt; Practice</i>	-0,005	-0,045	0,188	-0,365	0,297
<i>Paysage2 -&gt; Ravageur</i>	0,176	0,175	0,120	-0,031	0,359
<i>Paysage2 -&gt; Adventic</i>	0,017	0,017	0,122	-0,233	0,197
<i>Practice -&gt; Ravageur</i>	0,222	0,182	0,180	-0,186	0,499
<i>Practice -&gt; Adventic</i>	-0,487	-0,478	0,209	-0,748	0,097
<i>Practice -&gt; Costspec</i>	<b>0,229</b>	<b>0,247</b>	<b>0,159</b>	<b>0,016</b>	<b>0,451</b>
<i>Practice -&gt; Prodbrut</i>	0,103	0,071	0,131	-0,239	0,234
<i>Ravageur -&gt; Costspec</i>	<b>0,739</b>	<b>0,706</b>	<b>0,115</b>	<b>0,496</b>	<b>0,893</b>
<i>Ravageur -&gt; Prodbrut</i>	<b>-0,504</b>	<b>-0,482</b>	<b>0,098</b>	<b>-0,651</b>	<b>-0,255</b>
<i>Adventic -&gt; Costspec</i>	-0,096	-0,100	0,139	-0,400	0,156
<i>Adventic -&gt; Prodbrut</i>	<b>-0,519</b>	<b>-0,506</b>	<b>0,118</b>	<b>-0,640</b>	<b>-0,329</b>
<i>Costspec -&gt; Margeuha</i>	<b>-0,216</b>	<b>-0,204</b>	<b>0,079</b>	<b>-0,341</b>	<b>-0,015</b>
<i>Prodbrut -&gt; Margeuha</i>	<b>0,632</b>	<b>0,688</b>	<b>0,107</b>	<b>0,482</b>	<b>0,896</b>
<i>Subsisdys -&gt; Margeuha</i>	0,025	0,129	0,227	-0,390	0,447

*Lecture* : les valeurs en **bleu** (respectivement en **rouge**) indiquent que selon l'estimation amorcée (résultant d'un bootstrap) la valeur du lien structurel est significativement **positive**, respectivement **négative**.

Parmi les effets directs modérés, on distingue : i) ceux négatifs, des coûts spécifiques sur la marge brute ; ii) ceux positifs, de couverture du paysage (*Paysage1*) sur les pratiques et les adventices, des pratiques sur les ravageurs et les coûts spécifiques, ainsi que des subventions sur la marge brute. Les effets indirects modérés sont ceux négatifs des adventices sur la marge brute.

Parmi les effets directs faibles, notons : i) ceux négatifs des adventices sur les coûts spécifiques ; ii) ceux positifs, par ordre décroissant d'importance, de complexité du paysage (*Paysage2*) sur les ravageurs et des pratiques sur le produit. Les effets indirects faibles sont : i) ceux négatifs, de couverture du paysage sur le produit et la marge ; ii) ceux positifs, de couverture du paysage sur les coûts spécifiques et des pratiques sur le produit.

Parmi les effets directs très faibles, on relève : ceux positifs de complexité du paysage sur les coûts spécifiques et les pratiques ; ii) ceux négatifs des adventices sur les coûts spécifiques. Les effets indirects très faibles sont : i) ceux positifs de complexité et de couverture du paysage sur les ravageurs, et des pratiques sur la marge brute ; ii) ceux négatifs de complexité du paysage sur le produit et la marge.

Ainsi, les effets majeurs sont ceux positifs des ravageurs sur les coûts spécifiques, et du produit sur la marge. Les effets mineurs qui semblent avérés sont : i) ceux négatifs des pratiques sur les adventices et ceux des ravageurs et des adventices sur le produit, et indirectement sur la marge ; ii) ceux positifs des pratiques sur les coûts spécifiques.

L'utilisation du ré-échantillonnage aléatoire permet de valider ces premières constatations, en particulier au niveau des liens structurels du modèle interne (cf. tableau 4 et figure 9).

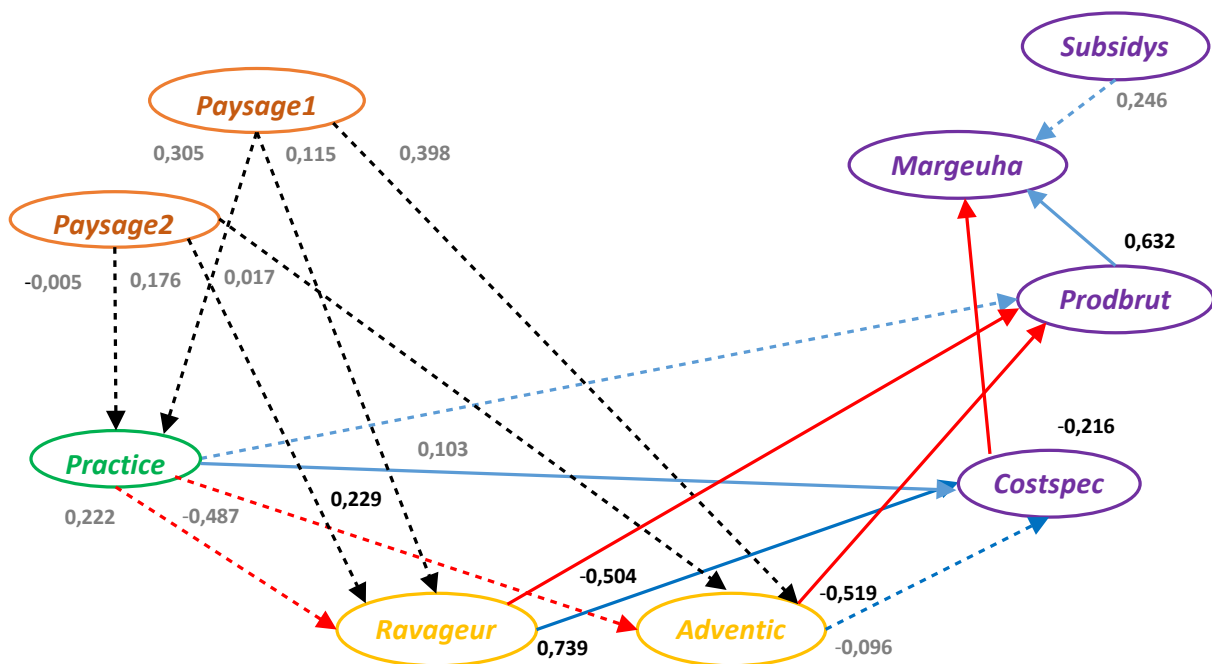


Figure 9. Schéma interne estimé du modèle structurel final après révision.

**Lecture :** les flèches en *bleu* (respectivement en *rouge*) indiquent que la valeur attendue du lien structurel est *positive*, respectivement *négative* ; les flèches en *noir* indiquent l'absence d'a priori sur la valeur attendue du lien structurel ; les valeurs (e.g., 0,398) indiquent la force du lien structurel reliant les variables latentes (**Paysage1** → **Adventic**) ; les liens structurels en pointillés (... >) ne sont pas significatifs selon l'estimation amorcée.

Ainsi, selon l'estimation amorcée avec 99 retirages aléatoires (*bootstrap*) et un niveau de confiance ( $\tau = 0,05$ ) à 95 %, les seuls liens significativement non nuls s'avèrent être :

i) d'une part, l'influence positive des pratiques sur les coûts spécifiques (**Practice** → **Costspec**  $\approx 0,25$ ), des ravageurs sur les coûts spécifiques (**Ravageurs** → **Costspec**  $\approx 0,71$ ) et du produit sur la marge brute (**Prodbrut** → **Margeuha**  $\approx 0,69$ ) ;

ii) d'autre part, l'influence négative des ravageurs sur la production (**Ravageur** → **Prodbrut**  $\approx -0,50$ ), des adventices sur la production (**Adventic** → **Prodbrut**  $\approx -0,48$ ) et des coûts spécifiques sur la marge (**Costspec** → **Margeuha**  $\approx -0,20$ ).

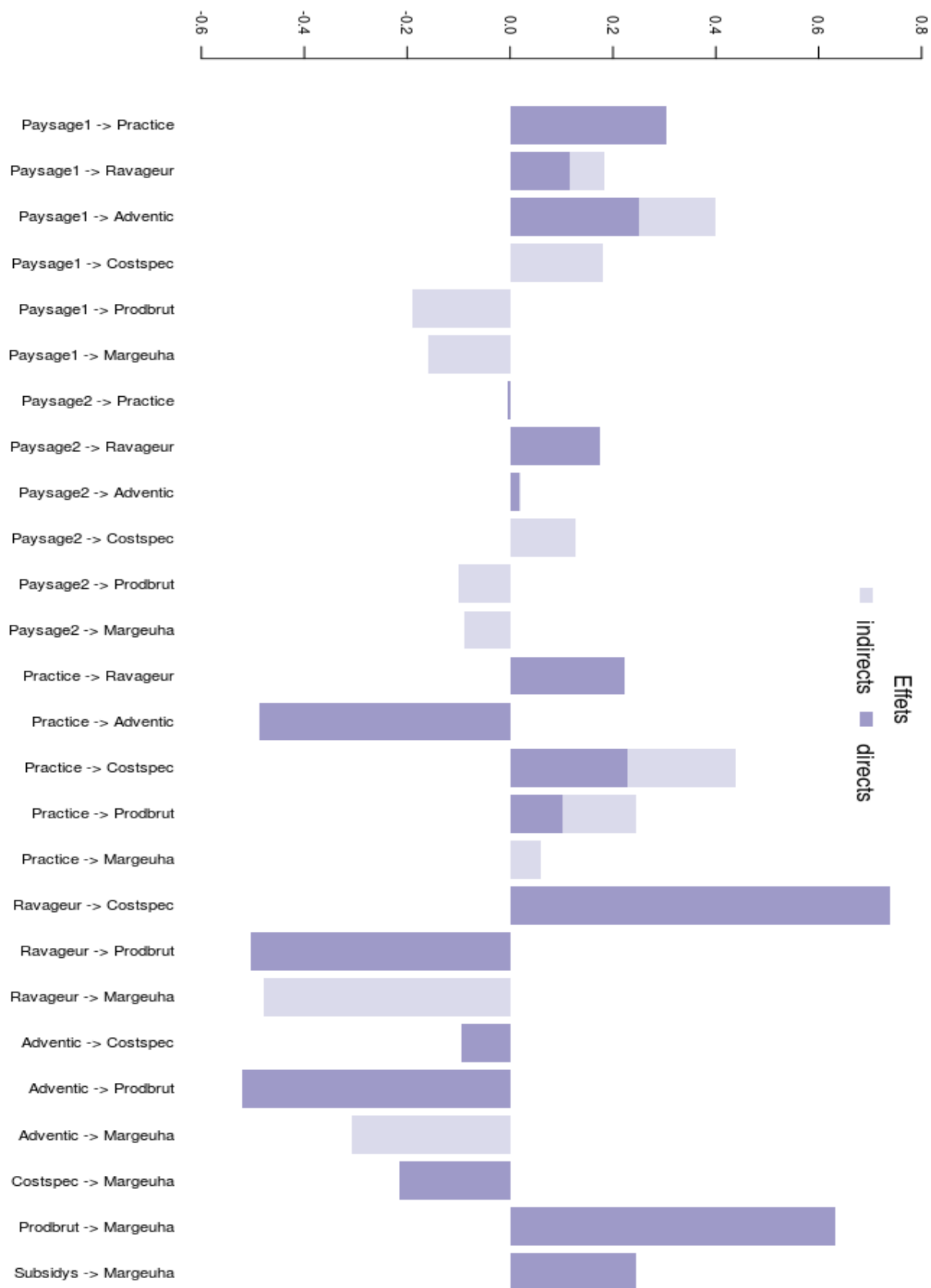


Figure 10. Distribution des effets directs et indirects

Lecture : les effets directs sont en bleu foncé et les effets indirects sont en bleu clair ; si effets directs et indirects co-existent, alors les bâtons sont empilés.

## Conclusion

L'objectif de ce travail est de proposer une introduction à une méthode alternative de modélisation structurelle adaptée à l'évaluation de pratiques écosystémiques, dans un contexte caractérisé par la complexité des interrelations entre paysages parcellaires, pratiques agronomiques et résultats économiques, ainsi que la faible taille des échantillons d'observations disponibles.

L'approche par les moindres carrés partiels a permis de mettre en évidence des effets statistiquement significatifs : i) négatifs sur les produits pour les adventices et les ravageurs; ii) positifs sur les coûts spécifiques pour les pratiques et les ravageurs ; et iii) de signe respectivement contraire sur la marge, négatifs pour les coûts spécifiques et positifs pour les produits.

Cependant, si les effets des pratiques agronomiques sur les adventices sont du signe attendu, elles ne le sont pas pour les ravageurs. Les signes des effets des paysages sur les ravageurs et sur les adventices ne sont pas non plus du signe attendu. Toutefois, ces valeurs estimées de ces effets ne sont pas statistiquement significatives.

Ainsi, il s'avère nécessaire de procéder à l'estimation de ce modèle structurel sur un échantillon plus important d'observations que permettraient l'extension de cette étude à une échelle spatiale plus large pour inclure d'autres paysages parcellaires et sa poursuite sur plusieurs années afin d'isoler d'éventuels effets de variations interannuelles.

Cet article est publié sous la licence Creative Commons (CC BY-SA).



<https://creativecommons.org/licenses/by-sa/4.0/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « Le Cahier des Techniques de l'Inra », la date de sa publication et son URL).

## Remerciements

L'auteur remercie les scientifiques correspondants de chacune des zones expérimentales (ZE) pour leur aide dans la collecte des données : Armin Bischoff pour la ZE Anjou-Saumurois, Claire Lavigne pour la ZE Basse Vallée de la Durance, Nicolas Parisey pour la ZE Armorique, et Sandrine Petit pour la ZE Côte d'Or.

## Bibliographie

- Amato S., Esposito Vinzi V., Tenenhaus M. (2004) A global goodness-of-fit index for PLS structural equation modeling. Communication au *Club PLS* (24/03), HEC Paris, France.
- Chin W.W. (1998) The partial least squares approach for structural equation modeling. In: Marcoulides, G.A. (Ed.), *Modern Methods for Business Research*. Lawrence Erlbaum Associates, Londres, pp. 295-336.
- Chin W.W., Newsted P.R. (1999) Structural Equation Modeling Analysis with Small Samples using Partial Least Squares. In : R. Hoyle (Eds.), *Statistical Strategies for Small Sample Research*. Sage Publication, Londres, pp. 307-341.
- Desbois D. (2018) *Livrable PEERLESS D1-2 : constitution et validation de la base de données agroécologiques et économiques*, Projet ANR-12-AGRO-006, 55 p.
- Efron B., Tibshirani R. J. (1993) *An introduction to the bootstrap*. Chapman & Hall, New York, 456 p.

## Le Cahier des Techniques de l'Inra 2020 (101)

- Franck P, Bischoff A., Bohan D., Klein E., Lavigne C., Martinet V., Memmah M., Parisey N., Petit S., Plantegenest M., Ricci B., Valantin-Morison M. (2017) Le projet PEERLESS : Viabilité d'une gestion écologique renforcée de la santé des plantes dans les paysages agricoles, In : *Ecologisation des systèmes de productions agricoles pour renforcer le contrôle biologique des bioagresseurs*, Franck et Petit (dir.), Inra, 27-28/11/2017, pp. 5-6.
- Hoyle R.H. ed. (1995) *Structural Equation Modeling: Concepts, Issues, and Applications*, Sage, Thousand Oaks, 290 p.
- Jonsson M., Straub C.S., Didham R.K., Buckley H.L., Case B.S., Hale R.J., Gratton C., Wratten S.D. (2015) Experimental evidence that the effectiveness of conservation biological control depends on landscape complexity. *Journal of Applied Ecology*, n° 52(5), 1274-1282.
- Lohmöller J.-B. (1989) *Latent Variable Path Modeling with Partial Least Squares*. Physica-Verlag, Mathematics, 283 p.
- Mezerette F. (2016) *Quels effets de la gestion agricole et du paysage sur l'abondance de bioagresseurs et le rendement?*, Mémoire du Master Ecologie Fonctionnelle, Comportementale et Evolutive, Université Rennes I, 34 p.
- Millennium Ecosystem Assessment (2005). *Ecosystems and human well-being : synthesis*. Washington, DC: Island Press.
- Petit S., Lavigne C., Ricci B., Aviron S., Plantegenest M., Tricault Y., Pollier A., Mézerette F., Ladet S., Biju-Duval L., Alignier A., Devienne S., Garambois N., Gibon A., Vialatte A., Choisit JP., Burnel L., Willm J. (2017) Les effets conjugués de la gestion parcellaire et du contexte paysager et de sa dynamique sur les bioagresseurs et les niveaux de régulation biologique, in *Colloque Ecologisation des Systèmes de Productions Agricoles*, Paris 27-28/11/2017.
- Sanchez G, Trinchera L., Russolillo G. (2017) plspm: tools for partial least squares path modeling (PLSPM). *Logiciel R*, version 0.4.9.
- Tenenhaus M., Esposito Vinzi V., Chatelin Y.M., Lauro C. (2005) PLS Path Modelling, *Computational Statistics & Data Analysis*, n°48, pp. 159-205.
- Wold H. (1981) *The Fix-Point Approach to Interdependent Systems*. Contributions to economic analysis 132, North-Holland, 335 p.

Tableau 5. Composition blocs thématiques correspondant aux variables latentes en variables manifestes selon leur nature

Variables latentes	Variables manifestes	Donnees		
		ecologiques	agronomiques	economiques
<i>Paysage1</i>	<i>Indicateurs de diversité paysagère</i>			
	Cultures d'hiver (% sole)	✓		
	Cultures de printemps (% sole)	✓		
	Cultures d'été (% sole)	✓		
	Cultures pérennes (% sole)	✓		
	Fabacées (% sole)	✓		
	Friche (% sole)	✓		
	Maraîchage (% sole)	✓		
	Oléagineux (% sole)	✓		
	Prairies (% sole)	✓		
Zones boisées (%sole)	✓			
<i>Paysage2</i>	<i>Indices de diversité paysagère</i>			
	Indice de Shannon grandes cultures	✓		
	Indice de Shannon toutes productions	✓		
<i>Practice</i>	<i>Pratiques agronomiques</i>			
	IFT total		✓	
	IFT autres		✓	
	IFT fongicides		✓	
	IFT herbicides		✓	
	IFT hors herbicides		✓	
	IFT insecticides		✓	
	IFT total hors traitement semencier		✓	
	IFT traitement semencier		✓	
	Nombre de passages		✓	
	Labour		✓	
	Ojectif de rendement de la fertilisation (t/ha)		✓	
	Fertilisation minérale en azote (u.N/ha)		✓	
	Fertilisation organique en azote (u.N/ha)		✓	
	Nombre de fertilisations minérales		✓	
	Nombre de fertilisations organiques		✓	
	Nombre de fertilisations		✓	
	Travail profond du sol (booléen)		✓	
	Travail superficiel du sol (booléen)		✓	
	Total travail du sol (h/ha)		✓	
<i>Ravageur</i>	<i>Ravageurs</i>			
	Sitobion avenae (puceron céréales)	✓		
	M.dirhodul (puceron céréales et rosier)	✓		
	Total pucerons	✓		
	Larves de Lema (criocères céréales)	✓		
	Dégats criocères (%)	✓		
<i>Adventic</i>	<i>Adventices</i>			
	Couverture totale par adventices (%)	✓		
	Dicotylédones (%)	✓		
	Graminées (%)	✓		
	Nombre d'espèces	✓		
	Pourcentage de surface en dicotylédones	✓		
Pourcentage de surface en graminées	✓			
<i>Costspec</i>	<i>Coûts spécifiques</i>			
	Coût semis (€/ha)			✓
	Coût préparation du sol (€/ha)			✓
	Coûts préparation et semis (€/ha)			✓
	Coûts désherbage chimique et mécanique (€/ha)			✓
	Coûts fongicides (€/ha)			✓
	Coûts insecticides et molluscides (€/ha)			✓
	Coûts phytosanitaires (€/ha)			✓
	Coûts fertilisants (€/ha)			✓
	Coûts de récolte (€/ha)			✓
Coûts spécifiques totaux (€/ha)			✓	
<i>Prodbrut</i>	<i>Produits bruts</i>			
	Rendements observés (t/ha)		✓	✓
	Prix (€/t)			✓
	Produit principal (€/ha)			✓
Produit total (€/ha)			✓	
<i>Subsidys</i>	<i>Subventions</i>			
	Droit à paiement unique (€/ha)			✓
	Aides couplées (€/ha)			✓
	Soutien à l'agriculture biologique(€/ha)			✓
Total des aides (€/ha)			✓	
<i>Margeuha</i>	<i>Marge unitaire à l'hectare</i>			
	Marge brute unitaire (€/ha)			✓
	Marge brute incluant les DPU (€/ha)			✓
	Marge brute standard (€/ha)			✓