Structural modelling: an application to the evaluation of ecosystem practices at the plot level

Dominique Desbois

Structural modelling: an application to the evaluation of ecosystem practices at the plot level

Dominique Desbois[1]

[1] UMR Economie publique, INRAE-AgroParisTech, Université Paris-Saclay 16 rue Claude Bernard, F-75231 Paris Cedex 05, France
(E-mail: dominique.desbois@inrae.fr)

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 by partial least squares. 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 implementation of the Partial Least Squares - Path Modeling (PLS-PM) approach. Finally, we conclude on the validation of this approach and the prosp[i]ects for its possible extension.
Keywords: ecosystem-based practices, structural modeling, partial least squares, plot scale, biological pest, specific costs, gross margin.

1 Introduction

The biological control of crop pests and weeds is an example of ecosystem services (ES) - the benefits that ecosystems provide to humankind (Millennium Ecosystem Assessment[11]) - provided at the farm plot level. The amenities provided by ecosystem infrastructures and agro-ecological practices in terms of soil protection, water resource management, and preservation of the habitat of agricultural auxiliaries constitute productive services that can be evaluated at the plot and farm level. For example, the biological control of crop pests and weeds by naturally occurring beneficials (such as ladybirds that predate aphids) is one of the productive services that can be mobilized at the plot level to reduce the dependence of agricultural production systems on pesticides. However, the relationships between crop management methods, levels of pest control by beneficials and crop yields are still insufficiently assessed (Franck *et al.*[6]). The structure of plot landscapes can also influence the level of pest abundance, so the relationship between plot landscape and yield remains to be more comprehensively inventoried (Jonsson *et al.*[8]). The objective of this work is therefore to analyze, via a Partial Least Squares - Path Modeling (PLS-PM) approach, the relationships between: i) plot landscapes, ii) crop pests and weeds, iii) agronomic practices, and iv) economic results.

The first part of this paper presents the specificity of the PLS-PM approach; the second part presents the field of study and describes the data considered in this work; the third part presents the results of the implementation of the PLS-PM approach; finally, the last part presents the conclusions on the validity of this approach and the perspectives of its application.

2 Structural equation modelling using partial least squares

Structural equation modelling mainly allows the study, via a hypothetical model specified in the form of equations, of the causal links (relationships) between several variables in order to account for the theoretical functioning of the system studied (Hoyle[7]). In this structural equation modelling, the variables can be either directly derived from observations or measurements (referred to as "manifest" variables) or not directly observable (referred to as "latent" variables). The PLS-PM approach is a variant of structural equation modelling that allows for the analysis of a complex system of relationships between the different variables under study, based on an a priori causal model (Path Modeling - PM) describing the relationships between the explanatory or 'exogenous' variables and the explained or 'endogenous' variables (Tenenhaus *et al.*[12]). The particularity of the PLS-PM approach lies in the fact that the estimation of the links of the structural model (path coefficients) is based on the Partial Least Squares (PLS) estimation criterion, rather than the Maximum Likelihood (ML) criterion, classically used in structural equation modelling.

The properties of Partial Least Squares (PLS) regression for estimating interdependent systems, established by Wold[13], led Lohmöller[9] to propose the PLS approach to structural equation modelling, PLS-Path Modelling (PLS-PM). Thus, the use of the PLS-PM approach does not require any assumptions on the distribution of variables (e.g. normality of the distribution) and is suitable for small sample sizes. Recent theoretical and algorithmic developments (Tenenhaus *et al.*[12]) have opened the field of its application more widely to multidisciplinary research where many groups of variables are likely to interact to condition social phenomena or economic behaviour. Indeed, such multidisciplinary research can often only be conducted for data sets where the conditions relating to the normality of the distribution, independence between observations, or sample size are not met (Chin and Newsteed[3], p. 314).

The specification of a PLS-PM model involves the following steps: i) specification of an initial hypothetical model describing the a priori relationships between the latent and manifest variables; ii) estimation of the model parameters via appropriate statistical software; iii) assessment of the goodness of fit of the structural model to the data (GoF); and iv) when the goodness of fit of the model is judged to be satisfactory, a final step is the interpretation of the results.

In the PLS-PM approach, the structural model is a set of conceptual constructs (or 'latent variables') linked by hypothetical causal relationships (the 'internal model') that can be estimated by means of measured or observed 'manifest variables' reflecting or, respectively, forming the latent variables (external model). Figure 1 illustrates the concepts of structural modelling, specifying the internal and external models and describing the relationships between latent and manifest variables.

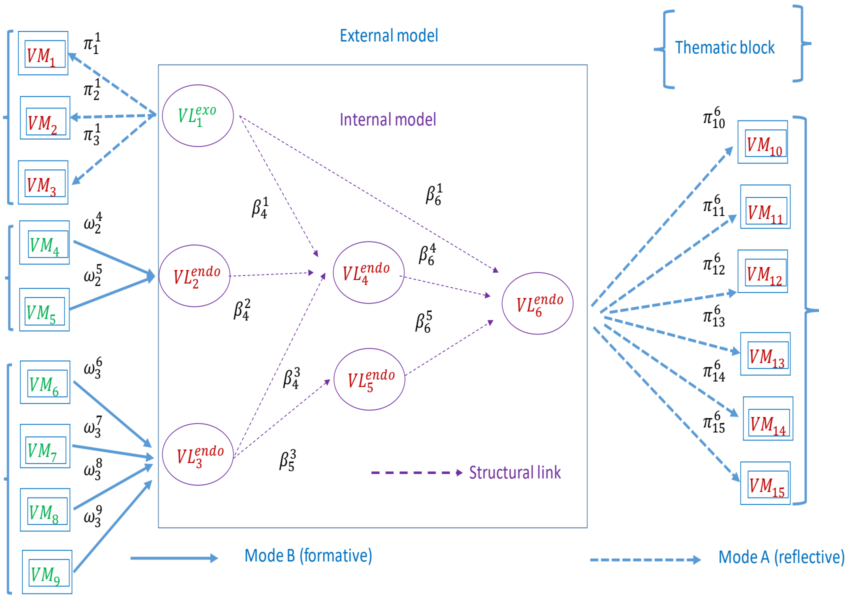

Fig. 1. Concepts of structural modelling

Reading: The beta coefficients represent the 'links' of the internal model; the manifest variables VM_h are associated with the exogenous VL_k^{exo} or endogenous VL_k^{endo} latent variables according to a mode that can be either 'reflective' (mode A): manifest variables 'reflect' latent variables, or 'formative' (mode B: manifest variables 'form' latent variables).

2.1 Specification of the internal model

The structural relationships between latent 'endogenous' (explained) and 'exogenous' (explanatory) variables constitute the internal model and are formalized by the following linear equations:

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

where β_l^k, called the 'structural link' (path coefficient), represents the sign and strength of the deterministic relationship between the endogenous latent variable LV_l^{endo} and the exogenous latent variables LV_k^{exo} . The part not explained by the deterministic model of LV_l is relegated to the residual ε_l. The structural links are estimated by a geometric projection (least squares) procedure whose only assumption is the independence between the deterministic part and the residual part, imposing that the covariance between each exogenous latent variable and the residual is zero ($cov(LV_k^{exo}, \varepsilon_l) = 0$).

2.2 Specification of the external model

Latent variables are defined by the "manifest variables" (VMs), derived from measurements and/or observations whether direct or indirect, via two modes (see Figure 1): i) the "reflective" mode where latent variables are reflected through their effects or consequences on the observed indicators; the "formative" mode where latent variables are supposed to be formed or constituted by the measured variables.

In the A reflective mode, the X_k manifest variable reflects the VL_k latent variable with mean *m* and standard deviation *1*, according to the following projective scheme (least squares regression):

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

where the π_h^k coefficient is the "outer weight" of the latent variable influencing the manifest variable. The residual $\varepsilon_h$has zero mean and is independent of the latent variable ($cov[VL_k, \varepsilon_h] = 0$).

In the B formative mode, the measured variables "form" the latent variables, according to the following equation:

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

where the ω_k^h coefficient is a "structural loading" contributing to the latent variable. The δ_k residual has zero mean and is independent of each of the manifest variables ($cov[X_h, \delta_k] = 0$).

The most commonly used mode is the A reflective mode. In estimating the model parameters, the PLS-PM approach aims to maximize the overall explained variance of the endogenous variables.

2.3 Validation statistics for the external model

The one-dimensionality of the block of manifest variables corresponding to each latent variable is a structural assumption of the external model that should be validated using the different criteria presented below.

i) Difference between the first two eigenvalues of the data block

First criterion of one-dimensionality, the principal component analysis of the block of data corresponding to each of the latent variables (see Table 1) provides a first criterion of one-dimensionality adapted from the Kaiser rule: if the first eigenvalue of the correlation matrix is greater than 1 and the second eigenvalue is much smaller, this means that the vast majority of the manifest variables are positively correlated with the first principal component.

ii) Cronbach's alpha

The second criterion of one-dimensionality is Cronbach's alpha, the ratio of the sum of the co-variances over the variance of the sum of the *H* manifest variables of the data block corresponding to a latent variable, i.e.:

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

Cronbach's α is widely used in reliability analyses with the following rule: if this ratio is greater than *0.7*, then the block can be considered unidimensional.

iii) Dillon-Goldstein rho

The last criterion used is the Dillon-Goldstein rho, the ratio of the variance of the latent variable to the variance of its block of manifest variables, estimable by

$$\hat{\rho} = \frac{[\sum_{h=1}^{H} corr(X_h, t_1)]^2}{[\sum_{h=1}^{H} corr(X_h, t_1)]^2 + \sum_{h=1}^{H}(1 - [corr(X_h, t_1)]^2)}$$

where t_1 is the first principal component of the thematic block of manifest variables.

If the estimate of the Dillon-Goldstein ρ is greater than 0.7, then the block is considered one-dimensional.

The Dillon-Goldstein ρ is considered a better criterion than Cronbach's α by Chin[2] because it is based on the structural factors of the internal model, rather than on the correlations between the manifest variables of the external model implicitly making the assumption that the manifest variables are a priori equivalent to each other in defining a latent variable (τ-equivalence assumption).

iv) Communality

The com_k 'communality' of the k^{th} thematic block indicates the extent to which the variability of the manifest variables of the k^{th} block is restored by the scores of the k^{th} latent variable. The 'commonality' of the kth thematic block is equal to the weighted sum of the squares of the correlations between the manifest variables and the Y_k reduced centered latent variable, *i.e.* :

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

2.4 Overall validation of structural modelling

i) The average redundancy index

In order to link the predictive performance of the external measurement model to the consistency of the internal model components, the redundancy index calculated for each endogenous thematic block measures the share of variability of the manifest variables related to the Y_h latent variables explaining the k^{th} endogenous latent variable, Y_k^{endo}, *i.e.*:

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

It interprets like an index of the capacity to predict the observed values of the k^{th} latent endogenous variable.

The average redundancy index, $\overline{Red}$, computed on the set of K^{endo} endogenous variables, i.e.:

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

Gives then a global index of the capacity to predict the observed values of the endogenous latent variables of the model.

ii) The Goodness of Fit

Proposed by Amato *et al.*[1], the goodness of fit (GoF) of the model is defined by the squared root of the product of the average 'communality' over the average R^2, *i.e.* :

$$GoF = \sqrt{\overline{com} \times \overline{R^2}}$$
$$= \sqrt{\frac{\sum_{k=1}^{K} \sum_{h=1}^{H_k} corr^2(X_k^h, Y_k)}{\sum_{k=1}^{K} H_k} \times \frac{\sum_{k=1}^{K^*} R^2 \left(Y_k^{endo}, Y_{h:Y_h \to Y_k^{endo}}\right)}{K^{endo}}}$$

where $\overline{com}$, the 'average communality' is the weighted mean of the communalities of each of the thematic blocks, *i.e.*:

$$\overline{com} = \frac{1}{\sum_{k:H_k>1} H_k} \sum_{k:H_k>1} H_k com_k$$

Because for each block, the thematic communalities are the means of square of the correlation coefficients, the average communality is the mean of the set of the square of the correlation coefficient between the latent variables and their manifest variables.

iii) The Bootstrap

As the PLS-PM approach is not based on distributional assumptions, the use of bootstrap-based validation procedures (Efron and Tibshirani[5]) becomes necessary in both an exploratory and confirmatory approach. Bootstrapped estimates are computed for external weights, factors, structural links, communality and redundancy indices, and overall goodness of fit. The principle of the bootstrap procedure is to randomly draw B new samples (usually *B<100*) of *N* observations (the 'seeds') into the initial sample of observations in order to obtain an estimate of the quantile function φ, reciprocal of the cumulative distribution function. For example, the bootstrapped values of the structural links are estimated on a B bootstrap basis using a Monte-Carlo procedure, yielding empirical confidence intervals estimated at the (1-τ) level of the quantile function of the bootstrapped communalities $\hat{\varphi}_{link}^{B}$, reciprocal of the cumulative distribution function, *i.e.* :

$$[\hat{\varphi}_{link}^{B}(\tau/2)\,; \hat{\varphi}_{link}^{B}(1-\tau/2)]$$

3 Material and Method

3.1 Agro-ecological context of the study

In this study, we apply the PLS-PM approach to the agro-ecological context defined by the plots of experimental or agricultural fields observed in the research framework constituted by four study zones (ZE) (Figure 3). These researches are federated by the multidisciplinary project 'Predictive Ecological Engineering for Landscape Ecosystem Services and Sustainability' (Peerless), and funded by the French National Research Agency ANR).

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

where β_l^k, called the 'structural link' (path coefficient), represents the sign and strength of the deterministic relationship between the endogenous latent variable LV_l^{endo} and the exogenous latent variables LV_k^{exo} . The part not explained by the deterministic model of LV_l is relegated into the residual ε_l. The structural links are estimated by a geometric projection (least squares) procedure whose only assumption is the independence between the deterministic part and the residual part, imposing that the covariance between each exogenous latent variable and the residual is zero ($cov(LV_k^{exo}, \varepsilon_l) = 0$).

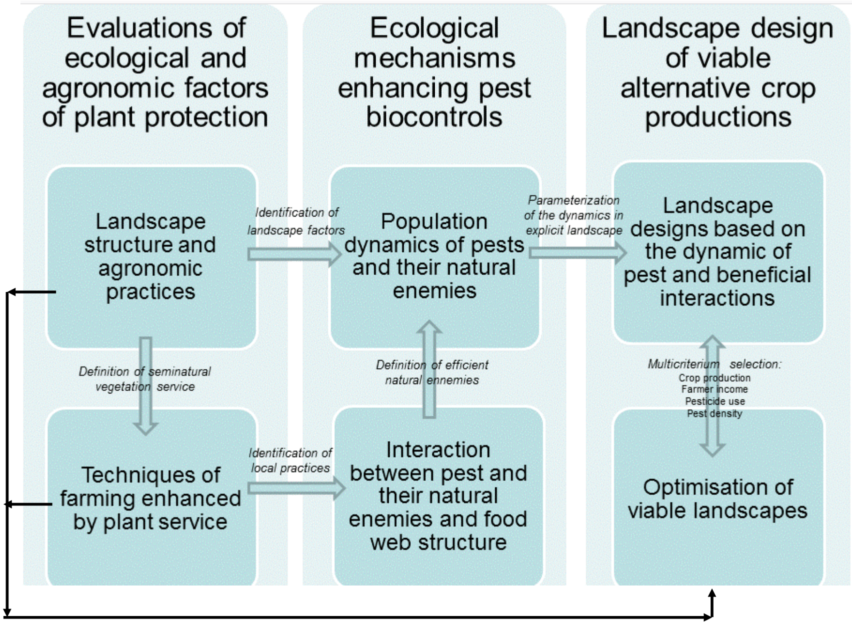

Fig. 2. Agro-ecological practices and infrastructure of the Peerless project

The Peerless project aims to identify the alternate managing strategies enhancing the pest control practices based on the functional biodiversity in arboriculture and in field crops to optimise the agricultural production systems, at the local landscape scales, in a double perspective of durability and economical viability of these productions (Franck *et al.*[6]). Peerless has been structured towards the three following objectives: i) the evaluation of agronomical and ecological factors of the plant protection; ii) the identification of ecological mecanisms enhancing the bio-control of pests; iii) the landscape conception of viable plant productions. the Peerless project aggregate four study zones (Figure 3) with field crops ('Anjou', 'Bittany', and 'Côte d'Or'), and one study zone with arboriculture ('Low Valley of the Durance'). These study zones (ZE) represent a set of 158 plots surveyed during 2014 and 2015.

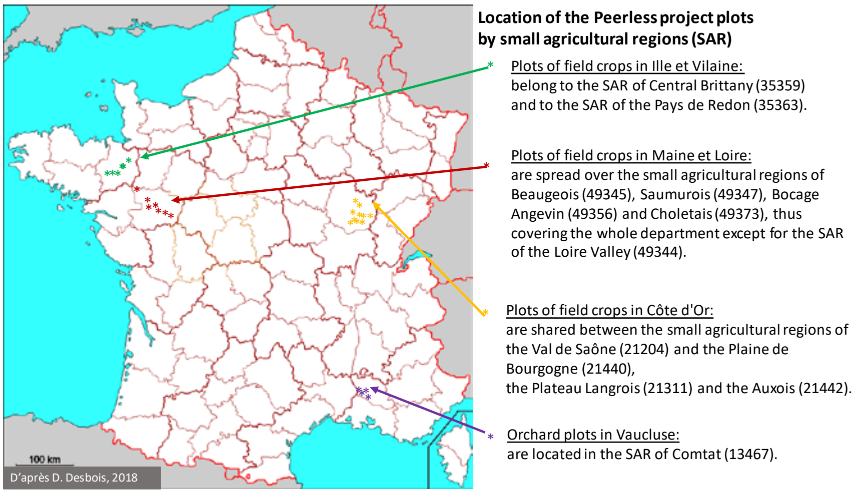

Fig. 3. Localization of the study zones in the Peerless project

3.2 Data

In this study, the Peerless project the data of considered at the scale of the different study zones are: i) agro-ecological measures from the field ; ii) field observations about the agronomical practices ; iii) economic estimates based on the agronomic practices, plot yields and the regional or national price references.

i) The agro-ecological data has been collected in the frame of the T1 task (landscape structure, field environment and agronomic practices) of the project. For each of the study zones, the landscape specification has been made within a circle of 1 km^2 of unit surface according two modalities. First specification of the landscape (Paysage1): the landscape contexts are specified by the percentages in winter crops, spring crops, summer crops, perennial crops, fabaceus, fallow, horticulture, seeds, and the ratios of meadows and of wooden areas. Second specification of the landscape (*Paysage2*), two Shannon indexes has been computed between distinct sites taking the lowest correspondence level of the common typologies: a) the first one computed on field crops; b) the second one computed on the whole set of the studied complexes (crops, meadows and wooden areas).

ii) The agronomic practices are documented by the indicators of cultural techniques, *i.e.*: the number of ploughs and soil works (deep and superficial); the number of fertilizations generally and particularly in nitrogen (N), as minerals as organics; the seedling and harvest dates; as well as the frequency treatment indicators (IFTs), as conventional as organic farming mode.

iii) The economic data have been produced in the framework of the ‘Optimization of viable landscapes’ task (T6) of the Peerless project and merged with the agro-ecological data (Desbois[4]) using together the price references issued from the French statistical office, those produced by the technical institutes (*Arvalis*, *Centre technique interprofessionnel des Fruits et Légumes*) and the agricultural offices (*FranceAgriMer*, *Chambres d'Agriculture*) to compute gross products, specific costs and gross margin.

3.3 The structural model and the estimation

Build from the concepts of the corporate accountancy, the simplified accounting relationship $Gross\_Margin = Gross\_Product - Specifics\_Costs$ defining the concept of gross margin as an algebraic sum of gross product and specific costs, offers a simple example of application not only to the farm holding but also to the cropped plots of the concept of internal structural model.

Once specified, this internal model allows to structure the set of measured or observed variables (the 'manifest' variables' - MVs) into several blocks corresponding to the conceptual artefacts, each block of manifest variables representing a latent variable. The particular composition into manifest variables of thematic blocks corresponding to different latent variables (*Landscape1*, *Landscape2*, *Pests*, *Weeds*, *SpecCost*, *GrossProd*, *Subsidies*, *Marginpua*) is given in annex (Table 5).

In a context of agro-ecological application, the concept of 'landscape' for the agricultural plot can be specified in a formative mode (B) by the occupation profiles of the soil (the different crops, the fabaceus, the fallow, the gardening, the seeds, the meadows and the wooden areas) as it can be in reflective mode (A) in the various indices de diversity that can be built from its description (« field crops » or « every production » Shannon indices).

Applied to eco-systemic contexts resulting from the conjunction of infrastructures and practices, the specification of a structural model of agro-ecological interaction and economic impact (Figure 5) allows to analyse the interrelations between ecological infrastructures (landscape profiles at the plot level), agronomic practices (ploughing, intermediary crops), and the induced results (growth of the product, reduction of the costs).

In this work, the initial structural model takes in the one proposed by Mezerette[10], completing it by the business management variables of costs, subsidies, product, and margin (Figure 1) and extending its application to the Peerless arboreal sites. This initial structural model specifies the *a priori* relationships between the agro-ecological variables describing the landscape (*Landscape1*, *Landscape2*), the agronomical practices (*Practice*) as well as the occurrence of diseases or pests (*Pests*) and weeds (*Weeds*), with the economic variables of specific costs (*SpecCost*), of gross production (*GrossProd*), of subsidies (*Subsidies*), and of gross margin per unit area (*Marginpua*) at the plot level.

In the scheme of the Figure 5, the ecological infrastructures (*Landscape1*, *Landscape2*) influence both the diseases or pests (*Pests*), the weeds (*Weeds*) and agronomical practices (*Practice*), in an a priori undetermined way (→). The agronomical practices impact a priori the gross products (*GrossProd*) and the specific costs (*SpecCost*) either directly by strengthening them (→), either indirectly by diminishing them (→) the pests and the weeds. The pests and the weeds are supposed to have a diminishing impact (→) on the gross products and expanding impact (→) on specific costs. The sum of these influences on the gross products and on the specific costs determines, with the subsidies (*Subsidies*), the impact in fine of the agro-ecological practices on the gross margin per hectare (*Marginpua*). In this scheme, the conceptual artefacts (*Landscape1*, *Landscape2*, and *Subsidies* appear as a priori exogenous (*i.e.* non determined by other phenomena) LVs while the concepts *Practice*, *Pests*, *Weeds*, *SpecCost*, *GrossProd*, and *Marginpua* are a priori endogenous (*i.e.* influenced by other phenomena) LVs.

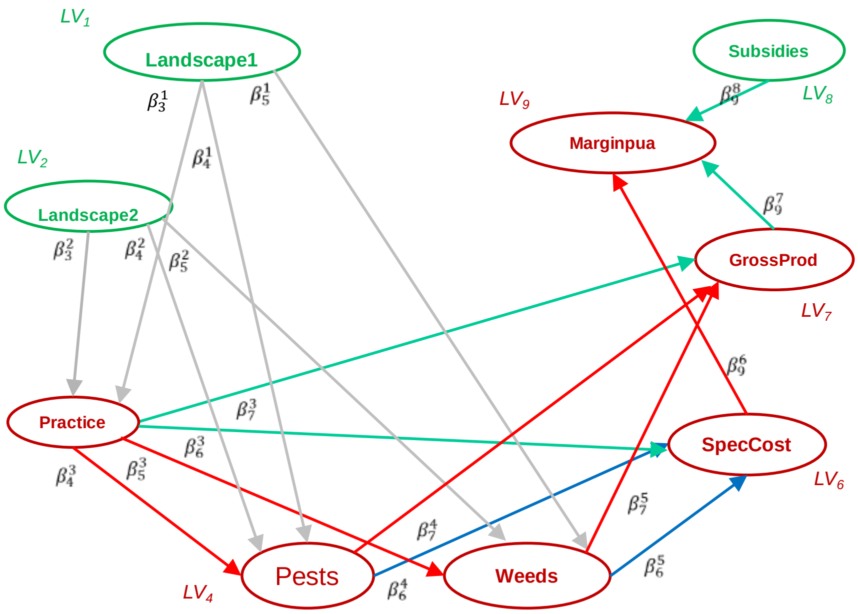

Fig. 4. The structural model of agro-ecological interaction and economic impact at the plot level

Reading of the graphical display: in green, the exogenous LVs; in burgundy, the endogenous LVs; in blue, the structural links of strengthening ; in red, the structural links of weakening ; in grey, the structural links a priori *undetermined.*

4 Results and discussion

The initial structural model is estimated on the basis of a set of 158 plots surveyed on 2014 and 2015. The setting of the PLS-PM approach has been realised via the R statistical software, version 3.6.2 (https://www.r-project.org/), and specifically with the 'pls-pm' package, version 0.4.9 (Sanchez, Trinchera and Russolillo, 2017). The estimation of structural links (path coefficients) of the model has been conducted using the reflective (A) mode for the set of latent variables.

Figure 6 displays the whole view of the external model. For example, for the latent variable *Pests*, the manifest variables reflecting this latent variable within the thematic block corresponding are: *S.avenae*, *M.dirhodum*, *Total.aphids*, *Lema.larvae*, *PctDamage.by.Lema*. The value of the structural factors is positive for the occurrences of pests (*Total.aphids*, *Lema.larvae*) and the index of their attacks (*PctDamage.by.Lema*). However, this value is negative for occurrences of the cereal ear aphid Sitobion (*S.avenae*), in contrast to the other pest occurrences of cereal and rose aphids (*M.dirhodum*) and cereal leaf beetles (*Lema*).

4.1 Checking the block one-dimensionality

The one-dimensionality statistics (Cronbach's alpha and Dillon-Goldstein's rho) computed on the basis of the initial structural model displayed (Figure 6) show too low values for the thematic blocks *Subsidies*, *Landscape2*, and *SpecCost* (Table 1). In fact, for these blocks, we observe some negatives correlations (loadings).

The blocks *Landscape2*, *Practice*, *Pests*, *Weeds* and *Marginpua* are considered as one-dimensional because they satisfy all the criterions. The block *Landscape1* and the blocks *SpecCost*, *GrossProd*, and *Subsidies* show values lower than 0.7 for the Cronbach's alpha, however the blocks *GrossProd* and *Subsidies* can be considered as unidimensional because they satisfy to the criterion of the Dillon-Goldstein's rho.

4.2 Fitting the external model and assessing the quality of the fit

The Goodness of Fit (*GoF*) value of the initial model is 0.4285. This index can be improved by adjusting the external model: indeed, some negatively correlated links can be transformed into positive links by appropriate recoding of manifest variables (MVs). Thus, in the block *Subsidies*, the recoding of the *CouplAid_pha* (coupled aid per hectare) MV into the *DecouplAid_ha = TotalAid_pha– CouplAid_pha* variable, makes it possible to obtain a positive correlation. Similarly, for the *Landscape1* thematic block, the *VegCrop_pct* (percentage of vegetable crop) MV is recoded into the *NonVegCrop_pct = 100 – VegCrop_pct* variable. For the block *Practice*, the *Fung_IFT* (index of the frequency of fungal treatments) MV is transformed into the *NonFung_IFT=Total_IFT – Fung_IFT* variable. For the block *SpecCost*, the *FungTCost_pha* (cost per hectare of fungal treatments) MV is transformed into the *NonFungTCost_pha =TotalCost_pha – FungTCost_pha* variable.

Blocks	Mode	Observed Var	Cronbach's Alpha	Dillon-Goldstein's Rho	Eigenvalue 1	Eigenvalue 2
Landscape1	A	9	0.398	0.073	**2.04**	1.67
Landscape2	A	10	0.473	0.642	**1.84**	1.45
Practice	A	18	**0.943**	**0.950**	**9.45**	**3.20**
Pests	A	5	**0.807**	**0.874**	**3.03**	**0.96**
Weeds	A	6	**0.871**	**0.904**	**3.68**	**1.09**
SpecCost	A	10	**0.834**	**0.879**	**4.75**	**2.30**
GrossProd	A	4	0.690	**0.805**	**2.30**	1.59
Subsidies	A	4	0.243	0.125	**2.09**	**0.18**
Marginpua	A	3	**0.986**	0.991	**2.92**	**0.07**

Tab. 1. One-dimensionality statistics of thematic blocks of the initial model.
Reading: the values in blue mean that the selected one-dimensionality criterion is satisfied.

After this recoding, the one-dimensionality statistics are enhanced (Table 2): they are all greater than 0.7 (acceptance level of the one-dimensionality hypothesis) except for the blocks *Landscape1*, *GrossProd*, and the *Subsidies* block with the values lower than 0.5 for the Cronbach's alpha, however near from the acceptance level with the Dillon-Goldstein's rho (>0,6) for *Landscape2*. The value of the Goodness of Fit for the amended model is enhanced in a marginal way, to 0.5078 (Table 3).

Blocks	Mode	Observed Var	Cronbach's Alpha	Dillon-Goldstein's Rho	Eigenvalue 1	Eigenvalue 2
Landscape1	A	10	0	0.215	**2.10**	1,79
Landscape2	A	2	**0.829**	**0.921**	**1.71**	**0.29**
Practice	A	19	**0.720**	**0.741**	**9.11**	**3.33**
Pests	A	5	**0.717**	**0.838**	**3.03**	**0.96**
Weeds	A	6	**0.871**	**0.904**	**3.68**	**1.09**
SpecCost	A	10	0.541	0.611	**3.02**	2.50
GrossProd	A	4	0.592	**0.790**	**2.30**	1.59
Subsidies	A	4	0.664	**0.804**	**2.08**	1.31
Marginpua	A	3	**0.986**	**0.991**	**2.92**	**0.07**

Tab. 2. One-dimensionality statistics for the thematic blocks of the amended model.
Reading: the values in blue mean that the selected one-dimensionality criterion is satisfied.

Theme block	Manifest variables	Communality	Type	R^2	Average redundancy
Landscape1	9	0.178	Exogeneous	0.0000	0.0000
Landscape2	2	0.850	Exogeneous	0.0000	0.0000
Practice	18	0.517	Endogeneous	0.0921	0.0477
Pests	5	0.606	Endogeneous	0.1254	0.0759
Weeds	6	0.608	Endogeneous	0.2806	0.1707
SpecCost	10	0.462	Endogeneous	0.6764	0.3127
GrossProd	4	0.542	Endogeneous	0.6906	0.3746
Subsidies	4	0.362	Exogeneous	0.0000	0.0000
Marginpua	3	0.972	Endogeneous	0.6391	0.6215
Weighted Mean		0.499			
Endogenous block Mean		0.618		0.417	0.2579
Goodness of Fit				0.5078	

Tab. 3. Communalities, R^2 and redundancies

Nota bene: the Goodness of Fit is computed from the endogenous blocks.

4.3 The structural model after revision

Taking in account the correlations between blocks result in the following structural model (Figure 4) where the relationships between latent variables are of expected sign except for the latent variable « Weeds » issued from the block *Weeds* displaying a negative relationship somewhat of low intensity, with the variable 'Specific Costs' issue du block *SpecCost.*

Take notice that the relationship between the latent variable *Practice* representing the agronomic practices and the latent variable *Pests* representing the occurrence of pest is positive while it is negative with the latent variable *Weeds* representing the presence of weeds. It would be interesting to distinguish between the specific practices of the pest control from those specific of the weed control.

The direct and indirect effects are displayed in Figure 6 as a categorical diagram, occasionally piled when they are cumulative. With regards to the direct or indirect effects, we distinguish those important (greater than 0.4), from those which are moderate (greater than 0.2 and lower than 0.4) and those which are low (lower than 0.2 and greater than 0.1), even from those which are very low (lower than 0.1).

Among the effects which are important, we have: i) the negative effects of the agronomic practices on the weeds; ii) the negative effects of the pest and the weeds on the gross product; iii) the positive effects of the gross product on the gross margin. The most important indirect effects (greater than 0.4) are those negative of the pest on the gross margin.

Among the moderate direct effects, we distinguish: i) the negative one, the specific costs on the gross margin; ii) the positive one, the cover of the landscape (*Landscape1*) of the practices on the weeds, of practices on the pests and the

specific costs, also of subsidies on the gross margin. The moderate indirect effects are the negative one of weeds on the gross margin.
Among the low direct effects, we note: i) the negative one of weeds on the specific costs; ii) the positive one, by decreasing importance order, of landscape complexity (*Landscape2*) on the pests and of practices on the gross product. The lower indirect effects are: i) the negative one, the landscape cover on the gross product and the margin; ii) the positive one, of landscape cover on the specifics costs and of practices on the product.
Among the lower direct effects, we notice: the positive ones of the landscape complexity on the specific costs and the practices; ii) the negative one of weeds on the specific costs. The very low indirect effects are: i) the positive one of landscape complexity and of the cover landscape on the pests, and of practices of the gross margin; ii) the negative ones of landscape complexity on the gross product and the gross margin.
Hence, the major effects are the ones positive of pests on the specific costs, and of gross product on the gross margin. The minor effects which appear as manifest are: i) the negative ones of practices on the weeds and the ones of pests and of the weeds on the gross product, and indirectly on the gross margin; ii) the positive ones of practices on the specific costs.
The use of random resampling validates these initial findings, particularly with regard to the structural links in the internal model (Table 4 and Figure 5).

Link	Estimate	Bootstrap Estimate	Standard-error	Q(0.025)	Q(0.975)
Landscape1 -> Practice	0.305	0.249	0.323	-0.521	0.639
Landscape1 -> Pests	0.115	0.203	0.251	-0.380	0.552
Landscape1 -> Weeds	0.398	0.431	0.215	-0.227	0.692
Landscape2 -> Practice	-0.005	-0.045	0.188	-0.365	0.297
Landscape2 -> Pests	0.176	0.175	0.120	-0.031	0.359
Landscape2 -> Weeds	0.017	0.017	0.122	-0.233	0.197
Practice -> Pests	0.222	0.182	0.180	-0.186	0.499
Practice -> Weeds	-0.487	-0.478	0.209	-0.748	0.097
Practice -> SpecCost	**0.229**	**0.247**	**0.159**	**0.016**	**0.451**
Practice -> GrossProd	0.103	0.071	0.131	-0.239	0.234
Pests -> SpecCost	**0.739**	**0.706**	**0.115**	**0.496**	**0.893**
Pests -> GrossProd	**-0.504**	**-0.482**	**0.098**	**-0.651**	**-0.255**
Weeds -> SpecCost	-0.096	-0.100	0.139	-0.400	0.156
Weeds -> GrossProd	**-0.519**	**-0.506**	**0.118**	**-0.640**	**-0.329**
SpecCost -> Marginpua	**-0.216**	**-0.204**	**0.079**	**-0.341**	**-0.015**
GrossProd -> Marginpua	**0.632**	**0.688**	**0.107**	**0.482**	**0.896**
Subsidies -> Marginpua	0.246	0.129	0.227	-0.390	0.447

Tab. 4. Bootstrapped estimation of the structural links of the internal model

Reading: values in blue (respectively in red) indicate that according to the bootstrap estimation the value of the structural link is significantly positive, respectively negative.

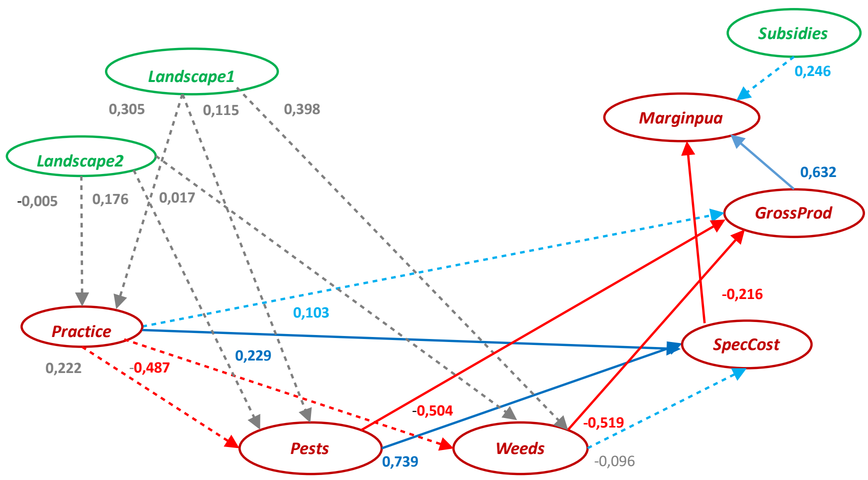

Fig. 5. Estimated internal scheme of the final structural model after revision

Reading: blue (respectively red) arrows indicate that the expected value of the structural link is positive, respectively negative; values (e.g., 0.398) indicate the strength of the structural link connecting the latent variables (e.g., Landscape1->Weeds); dotted structural links (···>) are not significant according to the bootstrapped estimation.

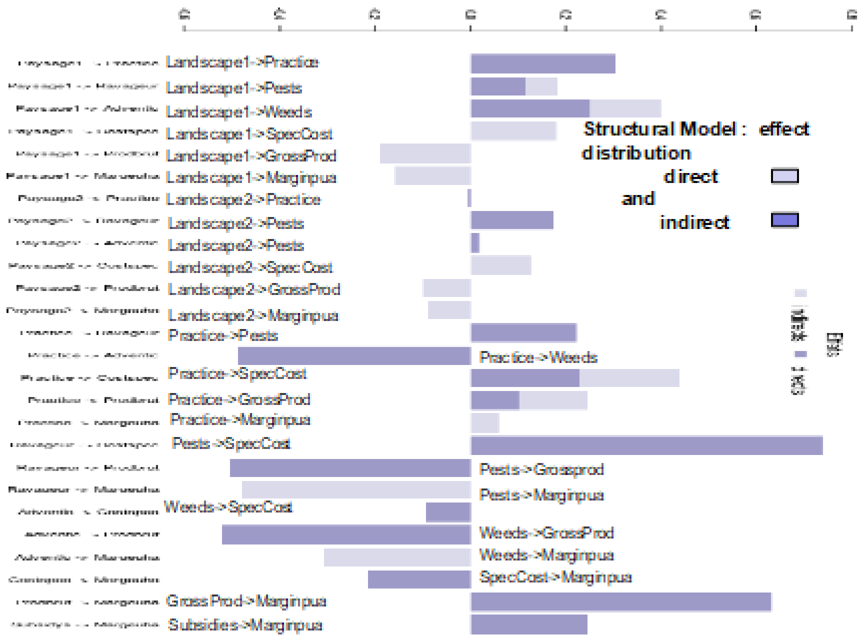

Fig. 6. Distribution of direct and indirect effects

Reading: direct effects are in **dark blue** and indirect effects are in **light blue**; if direct and indirect effects co-exist, then the bars are stacked.

Thus, according to the estimation initiated with 99 random removals (bootstrap) and a confidence level (τ=0.05) at 95%, the only significantly non-zero links turn out to be:

i) on the one hand, the positive influence of practices on specific costs (*Practice->SpecCost* ≈ 0.25), of pests on specific costs (*Pests->SpecCost* ≈ 0.71) and of product on gross margin (*GrossProd->Marginpua* ≈ 0.69);

ii) on the other hand, the negative influence of pests on production (*Pests->GrossProd* ≈ -0.50), weeds on production (*Weeds->GrossProd* ≈ -0.48) and specific costs on margin (*SpecCost->Marginpua* ≈ -0.20).

Conclusions

The objective of this work is to propose an introduction to an alternative structural modelling method adapted to the evaluation of ecosystem practices, in a context characterized by the complexity of the interrelationships between parcel landscapes, agronomic practices and economic results, as well as the small size of the available observation samples.

The partial least squares approach revealed statistically significant effects: i) negative on products for weeds and pests; ii) positive on specific costs for practices and pests; and iii) of opposite sign on margin, negative for specific costs and positive for products.

However, while the effects of agronomic practices on weeds are of the expected sign, they are not for pests. The signs of the effects of landscapes on pests and weeds are also not of the expected sign. However, these estimated values of these effects are not statistically significant.

Thus, it is necessary to estimate this structural model on a larger sample of observations that would allow the extension of this study to a larger spatial scale to include other plot landscapes and its continuation over several years in order to isolate possible effects of inter-annual variation.

References

1 Amato S., Esposito Vinzi V., Tenenhaus M. A global goodness-of-fit index for PLS structural equation modeling. Communication to PLS Club, HEC Paris, France, 2004.

2 Chin W.W. 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, 1998.

3 Chin W.W., Newsted P.R. 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, 1999.

4 Desbois D. Livrable PEERLESS D1-2 : constitution et validation de la base de données agroécologiques et économiques, ANR-12-AGRO-006 Project, 55 p., 2018.

5 Efron B., Tibshirani R. J. An introduction to the bootstrap. Chapman & Hall, New York, 456 p., 1993.

6 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. 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, Inra, pp. 5-6, 2017.
7 Hoyle R.H. ed. Structural Equation Modeling: Concepts, Issues, and Applications, Sage, Thousand Oaks, 290 p., 1995.
8 Jonsson M., Straub C.S., Didham R.K., Buckley H.L., Case B.S., Hale R.J., Gratton C., Wratten S.D. Experimental evidence that the effectiveness of conservation biological control depends on landscape complexity. Journal of Applied Ecology, n° 52, pp. 1274-1282, 2015.
9 Lohmöller J.-B. Latent Variable Path Modeling with Partial Least Squares. Physica-Verlag, Mathematics, 283 p., 1989.
10 Mezerette F. Quels effets de la gestion agricole et du paysage sur l'abondance de bioagresseurs et le rendement ?, Ecologie Fonctionnelle, Comportementale et Evolutive Master Memorandum, Rennes I University, 34 p., 2016.
11 Millennium Ecosystem Assessment. Ecosystems and human well-being: synthesis. Washington, DC: Island Press, 2005.
12 Tenenhaus M., Esposito Vinzi V., Chatelin Y.M., Lauro C. PLS Path Modelling, Computational Statistics & Data Analysis, n°48, pp. 159-205, 2005.
13 Wold H. The Fix-Point Approach to Interdependent Systems. Contributions to economic analysis 132, North-Holland, 335 p., 1981.

---

[i] This paper is dedicated to the memory of Jean-Paul Benzécri (1932-2019), emeritus professor at '*Université Pierre et Marie Curie*'.