Recovering cropping management practices specific production functions: clustering and latent approaches

Esther DEVILLIERS[1]* and Alain CARPENTIER[1]

[1] INRA, UMR SMART-LERECO,
4 allée Adolphe Bobierre - CS 61103,
35011 Rennes Cedex, France
*Corresponding author: esther.devilliers@inra.fr - +33(0)2.23.48.53.94

Contributed Paper prepared for presentation at the 93rd Annual Conference of the Agricultural Economics Society, University of Warwick, England - *Working version*

15 - 17 April 2019

Abstract. Reducing the use of pesticides and more generally of chemical inputs is a topical issue for governments. Economists generally advocate taxation for reducing polluting input uses. While econometric models tend to show that the pesticides price elasticity is low, these models mostly consider short-term adjustments. Mid-term adjustments of variable input uses are expected to be larger as reducing such uses require farmers to change their cropping management practices (CMPs). CMP is a notion closely related to the economists' production functions and used by agricultural scientists for characterizing crop production technologies. Yet, data lacking on farmers' CMPs prevents direct empirical analyses of CMPs' performances and adoption processes. The main objective of this paper is to propose original approaches for identifying farmers' CMPs in farm accountancy panel datasets with cost accounting. We consider that each CMP is characterized by a specific production function and propose approaches for identifying farmers' CMPs and the related production functions either sequentially or simultaneously. We demonstrate the relevance of our approaches through an empirical application based on a French arable crop farm accountancy unbalanced panel dataset covering the 1998-2014 period. Albeit preliminary, our empirical results demonstrate that our approaches perform relatively well. Indeed, they enable us to identify two wheat CMPs used by farmers: a low input CMP and a high yielding CMP.

Keywords: Cropping Management Practices · Production Function · Finite Mixture Models · Clustering Analysis.
JEL Code: Q12, Q16, C51.

Introduction

Regulation of pollutions due to the use of chemical inputs, of pesticides in particular, in agricultural production is a major policy objective in the European Union (EU).[1] However, most regulation policies implemented until now have achieved limited reductions in the use of these polluting inputs, especially regarding pesticides. These policies have been mainly based on market access restrictions, funding public research efforts on alternatives to chemical fertilization and crop protection as well as subsidies aimed to disseminate chemical input saving crop production practices. While economists generally advocate implementation of taxes for internalizing the negative external effects of inputs, public decision makers are reluctant to use taxes owing to their potential impact on farmers' income.[2] Crop production technologies being generally considered strongly dependent on chemical inputs, taxing these inputs is expected to significantly impact farmers' income with limited impact on their uses.

As a matter of fact, chemical inputs are key production factors in the crop production technologies used by farmers in industrialized countries (*e.g.*, Matson *et al*, 1997 [28]; Tilman *et al*, 2002 [42]; Lin, 2011 [22]), especially in the UE (*e.g.*, Aubertot *et al*, 2005 [2]). Production practices targeting high yield levels and based on short crop rotation schemes require high chemical fertilization levels. They generate high potential yield levels, which are worth protecting, and increase pest, disease and weed risks. As a result, these practices call for high protection levels that can be achieved by farmers, relatively easily and at reasonable (private) costs, by relying on chemical pesticides [3].

Econometric results generally demonstrate that farmers' pesticide uses display very limited responsiveness to pesticide price increases (see, *e.g.*, Skevas *et al*, 2013 [39]; Bocker *et al*, 2017 [7]), thereby providing support to the hypothesis stating current agricultural production technology "heavily depend" on pesticide uses. Yet, micro-econometric analyses of farmers' chemical input uses are generally based on panel data with short time dimension and assume that farmers' production technology remains unchanged throughout the considered period. Hence, they mostly reveal that farmers' chemical input choices are inelastic given their current technology choices.

These results are consistent with the view of agricultural scientists, according to which farmers cannot significantly modify their chemical input uses without changing

[1] Production and use of mineral nitrogen emit greenhouse gases and excess nitrogen uses pollute water bodies. Mineral phosphorus is a nonrenewable resource with rapidly decreasing stocks that generate eutrophication of surface when used in excess.

[2] For instance, pesticide taxes were considered by the European Commission but have not been implemented. In the few countries where they are, pesticide taxes are generally implemented with low tax rates and mostly for fund raising (Skevas *et al*, 2013 [39]; Finger *et al*, 2017 [14])

[3] Indeed, conventional crop production practices have been designed while taking for granted that high crop protection levels can be achieved at reasonable (private) cost with chemical pesticides (see, e.g., Aubertot *et al*, 2005 [2]). The economic efficiency of chemical crop protection also explains the focus of plant breeding on productivity rather than on resistances during the past decades (Vanloqueren and Baret, 2009 [43]).

their cropping management practices (CMPs) or cropping systems. Properties CMPs, which consist in ordered sequence of operations aimed to produce a given crop, are only guaranteed to hold in a limited range of chemical input uses (in a given production area). Adopting alternative production practices is necessary for farmers to significantly reduce their use of chemical inputs with limited impact on their income. Importantly, adopting a new CMP is an investment decision involving initial costs, adaption costs and expected return flows [4]. This implies that CMP adoption decisions are medium (to long) run decisions.

As CMP changes are necessary for significant changes in chemical input use levels by farmers, investigating famers' CMP choices is required for assessing the impacts of policy instruments aimed to reduce the use of these inputs. Unfortunately, the data sets usually used by agricultural production economists, which mostly come from farm accountancy data, don't contain relevant information for revealing farmers' CMP choices. These data document input purchases, crop yield levels and the related prices. But, they don't provide information on important components of CMPs such as seeding dates and densities, seed variety, tillage practices, *etc.*

Taking for granted this information lacking, the main objective of this article is to propose statistical and micro-econometric approaches for uncovering CMPs used by farmers from farm accountancy data (with cost accounting). Given that intensity in the use of chemical inputs of a CMP is directly related to the yield level targeted by this CMP, our methods aim to identify CMPs used by farmers based on their yield levels and chemical input uses.

Following Féménia and Letort (2016) [11], we take advantage of the conceptual similarity between agronomists' CMPs and economists' production functions. While CMPs describe in details how various inputs and cropping operations are combined to produce a given output, production functions focus on the role of purchased inputs. We assume here that a yield function corresponds to any CMP. This function describes how crop production levels respond to purchased input use levels. The functional form of this function, including its parameter value in a parametric setting, accounts for the unobserved components of the considered CMP.

We consider two approach types for uncovering CMPs in farm accountancy panel data. Both of them make use of farmers' series of input use and yield levels for identifying clusters of farmers using homogenous CMPs or production functions. First, we consider statistical nonparametric approaches, namely the k-means (kMC) clustering and ascending hierarchical classification (AHC) methods. These approaches are mostly descriptive and deliver farm clusters characterized by similar input use and yield levels. Second, we consider a micro-econometric modelling framework based on "latent pro-

[4] Initial investment costs may include new machinery. It always include information gathering. For instance, farmers may consult experts, neighbor farmers or various literatures. They may also monitor their fields for assessing the performances of the new practices and, then, for adapting it to the context of their farm.

duction technologies". We specify a structured micro-econometric model assuming that farmers can use either of a given number of production technologies corresponding to distinct CMPs. This model is defined as a latent class model, that is to say a parametric mixed model with a discrete mixing probability distribution (McLachlan and Peel, 2000 [30]; McLachlan and Krishnan, 2007 [29]). Estimating this model then yields the parameters of the underlying production technologies and the share of farmers using each of these technologies. The technology or CMP that is the most likely used by a given farmer can then be determined from a probabilistic viewpoint.

We illustrate our approaches by investigating French farmers' CMPs for winter wheat based on a panel dataset from 1998 to 2014 covering *la Marne crayeuse*, a small homogenous (highly productive) arable crop production area located in eastern France.[5] We consider two CMP types, "high yielding" (HY) CMPs and "low input" (LI) CMPs. HY-CMPs are considered as the conventional CMPs in France. They are designed to achieve high yield levels, near the maximum potential production level given the considered agro-climatic conditions. LI-CMPs are designed as chemical input, especially pesticide, saving CMPs (Meynard 1985 [31], 1991 [32]). As LI-CMPs aim to achieve yield levels (slightly) lower than those targeted by their corresponding HY versions, LI-CMPs are expected to be more profitable that HY-CMPs when wheat price is sufficiently low and/or chemical inputs are sufficiently expensive. These CMPs were developed by INRA starting in the early 80s and promoted starting in the mid 90s (Loyce *et al*, 2008 [25]; Loyce *et al*, 2012 [24]; Larédo and Hocdé, 2014 [21]). *La Marne crayeuse* is of special interest in this context because it is among the first area where LI-CMPs were tested through on-farm experiments. Moreover, these experiments were supported by local stakeholders, including extension services and down-stream industries (Loyce and Meynard, 1997 [23]).

Although still preliminary, our empirical results demonstrate that the proposed statistical and micro-econometric approaches allow identification of CMPs characterized by their target yield levels, or equivalently, their chemical input use intensity levels. In particular, when considering two CMPs, our estimation results systematically display clusters of farmers with expected characteristics (on average per cluster), namely a "LI cluster" and a "HY cluster". Farmers in the LI cluster achieve lower yield levels and use larger amounts of chemical inputs that farmers in the HY cluster. Importantly, the average yield and input use levels characterizing the LI and HY clusters conform to those characterizing the LI and HY CMPs considered by Loyce and Meynard (1997) [23], Loyce *et al* (2008) [25] and Loyce *et al* (2012) [24]. Moreover, the trends in the adoption rate of the LI-CMPs uncovered by our approaches conform to the evolution of wheat and input prices observed in the data, at least according to standard production economics principles.

The pioneering work of Griliches (1957) [17] originated an abundant literature on the adoption of agricultural production technologies, in economically developed or developing countries. Micro-econometric studies rely on data in which the adoption decisions of

[5] *La Marne crayeuse* is located in the *Champagne* region, about 150 km from Paris.

the considered techniques are observed. They generally focus on the adoption of specific techniques or practices (*e.g.*, use of a cultivar, tillage techniques, integrated pest management) and put emphasis on specific determinants such as learning processes (*e.g.*, Foster and Rosenzweig, 2010 [15]; Sadoulet, 2016 [5]), heterogeneity in return to adoption (*e.g.*, Suri, 2011 [40]; Michler *et al*, 2018 [33]) or labour constraints (*e.g.*, Fernandez-Cornejo, 2005 [13]). Specific analyses aim to assess the impacts of new technologies on yields, input uses and income (*e.g.*, Fernandez-Cornejo, 1996 [12]; Khanna, 2001 [20]; Teklewold *et al*; 2013 [41]). Our use of "latent technologies" relates our modelling framework to that used in the strand of studies originated by the work of Orea and Kumbhakar (2004) [35] and Greene (2005) [16] on latent class stochastic frontier models (*e.g.*, Alvarez and del Corral, 2010 [1]).

To our knowledge, our study is the first investigating the effects and drivers of technology adoption while simultaneously recovering the adoption decision. Our modelling approaches are of special interest for two main reasons. First, production practices are rarely documented in datasets typically used by agricultural production economists. Second, considering production practice changes is crucial when assessing the long run impacts of agri-environmental policies.

The rest of article is organized as follows. The first section discusses LI-CMPs, their history and their underlying agronomic principles. The second section presents a simple micro-economic modelling framework aimed to highlight the impacts of CMP choices on farmers' input demands and yield supplies. The third section presents our empirical analysis. First, we briefly present the data, the statistical and micro-econometric approaches used for uncovering farmers' CMP choices and their characteristics in terms of input use and yield levels. Then, we present the results related to the adoption of LI-CMPs by wheat growers in the *la Marne crayeuse* area.

1 Agronomic principles and brief history of "Low Input" CMPs

Cropping management practices (CMPs) are defined by agronomists as ordered sequence of yield production decisions or decision rules aimed to produce a given crop. CMPs include soil preparation operations and type, seeding type, date and density, fertilization and pesticide applications, etc. We are interested in specific CMPs, the LI- CMPs proposed by agronomists in the late 80s and, then, developed and promoted by agronomists and extension agents since the mid 90s. LI-CMPs were developed by INRA starting in mid 80s and combined with multi-resistant wheat cultivars in the late 90s. The "LI-CMP and hardy wheat cultivars" package were promoted by agronomists, extension agents and French wheat breeders starting in the late 90s (Larédo and Hocdé, 2014 [21]).

1.1 "Low Input" CMPs as induced technological innovations

LI-CMPs can be interpreted as technological innovations aimed to provide answers to two main issues raised by conventional HY-CMPs. First, HY-CMPs are intensive in chemical input uses, which are polluting inputs. LI-CMPs were primarily designed for reducing pesticide uses. Second, the decrease in grain prices induced by the – progressive for cereals while sudden for oilseeds – removal of the CAP price support in 1992 called into question the profitability of grain production in the EU from the late 90s to the mid 00s. Due to the low grain prices during this period, HY-CMPs appeared to much less profitable than they were in the early 00s.

The price support implemented by the CAP until the so-called McSharry reform in 1992 led most agricultural scientists to develop HY-CMPs to be adopted by European grain producers. Indeed, due to the relative scarcity of arable land in Western Europe, adopting HY-CMPs appeared to be the most profitable technological option for farmers to benefit from high grain prices (Mahé and Rainelli, 1987 [26]; Meynard, 1991 [32]). HY-CMPs aim to increase grain potential yield by increasing seeding densities, choosing early seeding dates, relying on productive seed varieties and applying large amounts of, especially nitrogen, fertilizers. Importantly, these HY techniques tend to increase pest and weed pressures and, consequently, calls for efficient crop protection. Early seeding dates tend to expose crops to pest outbreaks. Nitrogen fertilizer use tend to trigger competition by weeds. High seed densities, productive – but susceptible to diseases – cultivars and high loads of nitrogen fertilizer tend to increase wheat susceptibility to diseases. Yet, availability of efficient, as well as relatively cheap, chemical pesticides enables farmer to control the pest and weed pressures triggered by HY techniques.

LI-CMPs were conceived by agricultural scientists as an agronomic response to the polluting emissions induced by the use of chemical inputs, of chemical pesticides in particular and to the decrease in grain prices due to the CAP reform implemented in 1992 – the so-called McSharry reform. HY-CMPs are conceived to achieve high target yield levels but rely on high levels of chemical input uses, precisely because the techniques implemented for achieving high target yield levels tend to trigger the need of high fertilization and crop protection levels. The basic principle of the conception of LI-CMPs is to lower target yield levels in order to lower chemical input uses, pesticides in particular. Lowering target yield levels directly reduces crop nutrition needs and, thereby, nitrogen fertilization uses. LI-CMPs lower crop protection needs by avoiding cropping techniques that increase pest and weed pressures. Therefore, they allow reducing pesticide uses.

The HY-CMPs and LI-CMPs considered by agronomists vary across time and production areas, depending on economic and agro-climatic conditions (Loyce and Meynard 1998 [23], Rolland *et al*, 2003 [38]; Bouchard *et al*, 2008 [8]; Loyce *et al*, 2008 [25]; Loyce *et al*, 2012 [24]). On average, the yield levels obtained with LI-CMPs are 10% lower than those obtained with HY-CMPs. Nitrogen fertilizer loads decrease by 10% from the HY-CMPs to the LI-CMPs while the use of (mostly) fungicides and insecticides is reduced by around 30%. Finally, due to the lower sowing densities in LI-CMPs seed uses decrease by around 50% when using these CMPs. Also, hardy wheat cultivars are complementary to the agronomic principles underlying the design of LI-CMPs (Loyce *et al*, 2008 [25]; Larédo and Hocdé, 2014 [21]). These cultivars are resistant to multiple diseases but slightly less productive than the ones typically used in HY-CMPs. Finally, LI-CMPs are labor and fuel saving thanks to their lower expected pesticide application numbers.

Unfortunately, no data exist on the adoption of LI-CMPs by French farmers. Moreover, farm accountancy data, even with cost accounting, don't contain any indicator enabling us to identify farmers using LI-CMPs. For instance, seed cultivars are not reported. Similarly, if purchased seed expenditures vary with sowing densities, these expenditures may also vary with seed prices and the share of seeds produced farmers themselves. This information lacking has two main, related, implications. First, we can only consider inferring farmers' CMPs from their yield and chemical input use levels, that is to say indirectly. This point is discussed in the last section. Second, the diffusion process of LI-CMPs among French arable crop producers cannot be tracked. Yet, it is possible to track the economic conditions that were more or less favorable to the adoption of LI-CMPs by farmers.

1.2 Recent trends in input and wheat prices, and their potential impacts on "Low Input" CMPs

Due to the CAP price support, the wheat price paid to French farmers was relatively high until the McSharry reform 1992 (see figure 1). Then it progressively declined and attained the world price level in the late 90's. World wheat price remained low until 2006. It dramatically increased in 2007, following the US biofuel policy and increasing fuel and feed demands. Wheat price has remained relatively high on average since then, though it is volatile due to random shocks on cereal supply and demand.

The early 90's are considered as the peak of the chemical intensity of the French arable production, with price ratios favoring the use of chemical inputs and, as a result, of HY-CMPs. LI-CMPs were developed starting in the mid 80s. They started to be tested on farm fields and promoted in the 90s, as responses to the low wheat prices observed from the late 90s to 2006. The yield reductions entailed in adoption of LI-CMPs have much less detrimental effects on profits when crop prices are low. But, the high wheat price levels observed since 2007 suggest that HY-CMPs might be as profitable today as they were in the early 90s. Yet, input prices increased since the early 90s, at difference rates though.

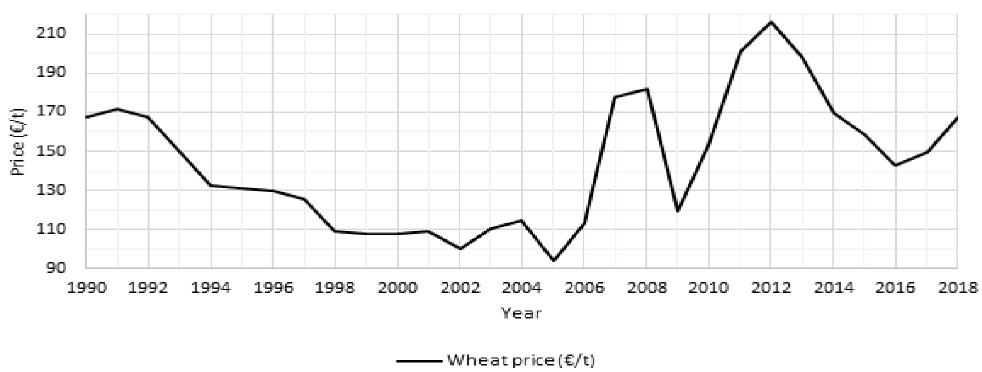
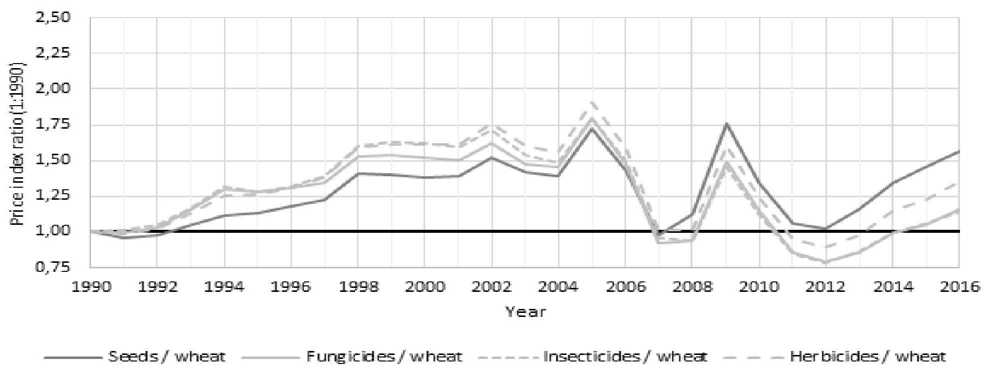


Fig. 1: Average farmgate wheat nominal price, France, 1990-2018

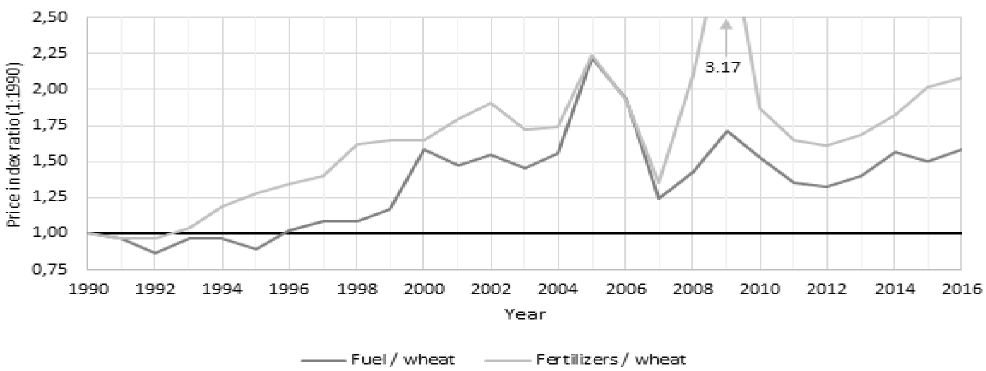
Figure 2 displays the evolution of the ratio of the input price indices to that of wheat from 1990 to 2016, using 1990 as the baseline. Basically, an input is cheaper (more expensive) for wheat production in year t than it was in 1990 when ratio of the price index of this input to that of wheat is below (above) 1.

Panel (a) in Figure 2 shows that pesticide prices were about 50% more expensive relative to wheat is the early 00s than they were in the early 90s. Since 2007, the ratios of the prices of pesticides to that of wheat have been only slightly above the ones observed in the early 90s. Indeed, the nominal prices of fungicides and insecticides remained steady while that of herbicides increased only by 15% from 1990 to 2016.

Fuel and fertilizers increased significantly during the 90s, as shown by panel (b) in Figure 2. Fertilizers were 75% more expensive relative to wheat during the 00s and 10s than they were during the early 90s. The corresponding ratio is around 50% for fuels. Differences in input uses between LI-CMPs and HY-CMPs are less important for fuel and fertilizer than they are for fungicides and insecticides. Yet, fuel and fertilizer price levels were substantially higher from 2007 to 2016 than in the early 90s.



(a) Seeds and pesticides



(b) Fuel and fertilizers

Fig. 2: Price index ratios, inputs to wheat, France, 1990-2016 (1 in 1990)

This factual analysis suggests that economic conditions tended to favor adoption of LI-CMPs from the late 90s to 2006, mostly due to the low grain prices observed during this period. The prices of fungicides and insecticides, the use of which is aimed to be reduced by adopting LI-CMPs, remained stable from 1990 to 2016. The high grain price levels observed since 2007 have tended to favor conventional HY-CMPs, although these effects of high grain prices on the profitability of HY-CMPs are partially offset by the high levels of fuel and fertilizer prices.

2 Cropping management practices, yield levels and input use levels

This section presents a simple micro-economic modelling framework designed to address the issues raised by accounting for CMPs in micro-economic models of agricultural production choice. We explicitly relate the concept of CMP to that of production function and then define define CMP choices as production function – or technology – adoption decisions. This enables us to analyze the role of CMPs on farmers' long run production decisions, with a particular focus on chemical input uses.

2.1 CMPs and production functions

Agronomists' CMPs are closely related to economists' production functions. The former is defined as an ordered sequence of cropping operations and input uses while the later focus on input uses. Let consider the following quadratic yield function:

$$y = f^c(\boldsymbol{x}) = b^c - (1/2) \times (\boldsymbol{d}^c - \boldsymbol{x})'\boldsymbol{A}^c(\boldsymbol{d}^c - \boldsymbol{x}) \text{ with } b > 0 \text{ and } \boldsymbol{d} > \boldsymbol{0} \tag{1}$$

for representing how chemical input uses $\boldsymbol{x} = (x_k : k \in K)$ lead to crop yield level y given CMP c. Yield function $f^c(\boldsymbol{x})$ is increasing in $\boldsymbol{x}$ as long as condition $\boldsymbol{d}^c \geq \boldsymbol{x}$ holds, where $\boldsymbol{d}^c = (d_k^c : k \in K)$.[6] Parameter matrix $\boldsymbol{A}^c = [a_{km}^c : (k,m) \in K \times K]$ is assumed to be positive definite, thereby ensuring the strict concavity of $f^c(\boldsymbol{x})$ in $\boldsymbol{x}$.

Quadratic yield functions are of interest here as second order approximations of any smooth yield function. Furthermore, as shown below their congruent yield input demand, output supply and profit function are obtained in analytical closed form. The specific parametrization of the yield function given in equation (1) was chosen for facilitating its parameter interpretation (Carpentier and Letort, 2012 [9]; 2014 [10]). Term b^c denotes the maximum yield level – hereafter denoted as the potential yield level – that can be achieved by using the production technology described by $f^c(\boldsymbol{x})$. Term $\boldsymbol{d}^c$ denotes the chemical input uses needed to achieve yield level b^c.[7] Matrix $\boldsymbol{A}^c$ characterizes the input effects on crop yield and their interactions.[8]

It is easily shown that quadratic production function (1) yields input demand, yield supply and profit functions in analytical closed forms (*e.g.*, Carpentier and Letort, 2012

[6] Yield function $f^c(\boldsymbol{x})$ can be generalized as a quadratic function with plateaus by replacing differences $b_k^c - x_k$ by truncated differences $\max\{d_k^c - x_k, 0\}$

[7] The assumed properties of $f^c(\boldsymbol{x})$ ensure that $b^c = \max_{\boldsymbol{x} \geq \boldsymbol{0}} f^c(\boldsymbol{x})$ and $\boldsymbol{d}^c = \arg\max_{\boldsymbol{x} \geq \boldsymbol{0}} f^c(\boldsymbol{x})$

[8] The smaller a_{kk}^c, the higher the marginal productivity of input k at low use levels of this input. Inputs k and m are "cooperant" in the sense of Rader (1968) [36] if and only if $a_{kk}^c \leq 0$. Increasing the use of input k (resp. m) cannot decrease the marginal productivity of input m (resp. k). Indeed, inputs are expected to be cooperant when they fulfil distinct roles in the considered production process (as is the case for inputs such as herbicides, insecticides, fungicides and nitrogen fertilizers). Function $f^c(\boldsymbol{x})$ represents a "normal" technology in the sense of Rader (1968) [36] if and only if all inputs are cooperant (*i.e.*, $a_{kk}^c \leq 0$ for any pair of distinct inputs).

[9]; 2014 [10]). Ignoring, for simplicity, price and production risks, crop price is denoted by p while vector $\boldsymbol{w} = (w_k : k \in K)$ collects chemical input prices. Solving farmers' gross return maximization problem, $\max_{\boldsymbol{x}\geq \boldsymbol{0}}\{pf^c(\boldsymbol{x}) - \boldsymbol{w}'\boldsymbol{x}\}$, yields the following input demand:

$$\boldsymbol{x}^c(p,\boldsymbol{w}) = \boldsymbol{d}^c - \boldsymbol{\Lambda}^c\boldsymbol{w}p^{-1}, \tag{2}$$

yield supply:

$$y^c(p,\boldsymbol{w}) = b^c - (1/2)\times\boldsymbol{w}'\boldsymbol{\Lambda}^c\boldsymbol{w}p^{-2}, \tag{3}$$

and gross return:

$$\Pi^c(p,\boldsymbol{w}) = pb^c - \boldsymbol{w}'\boldsymbol{d}^c + (1/2)\times\boldsymbol{w}'(\boldsymbol{\Lambda}^c)^{-1}\boldsymbol{w}p^{-1} \tag{4}$$

functions, where $\boldsymbol{\Lambda}^c = [\gamma^c_{km} : (k,m) \in K\times K] = (\boldsymbol{A}^c)^{-1}$. Of course, functions $y^c(p,\boldsymbol{w})$ and $x^c_k(p,\boldsymbol{w})$ increase in crop price p and decrease in input k price w_k.[9] Throughout this section we assume, for simplicity, that input yield use levels, $y^c(p,\boldsymbol{w})$ and $\boldsymbol{x}^c(p,\boldsymbol{w})$, are strictly positive.

2.2 "Low Input" versus "High-Yielding" CMPs

While the generic quadratic functional form of $f(\boldsymbol{x}; b^c, \boldsymbol{d}^c, \boldsymbol{A}^c)$ define how chemical input uses are combined to yield a given level of crop output, values of parameters $(b^c, \boldsymbol{d}^c, \boldsymbol{A}^c)$ depend on the CMP described by this yield function. Let $(b^h, \boldsymbol{d}^h, \boldsymbol{A}^h)$ denote the value of parameter $(b^c, \boldsymbol{d}^c, \boldsymbol{A}^c)$ corresponding to HY-CMPs and $(b^l, \boldsymbol{d}^l, \boldsymbol{A}^l)$ that corresponding to LI-CMPs.

Given that LI-CMPs are designed for achieving lower yield levels by relying on lower input use levels, we can assume that $b^l < b^h$ and $\boldsymbol{d}^l < \boldsymbol{d}^h$. We will also assume, for simplicity, that $\boldsymbol{A}^l = \boldsymbol{A}^h$.[10] Yield functions $f^h(\boldsymbol{x})$ and $f^l(\boldsymbol{x})$ are depicted in figure 3. Due to their conception, HY-CMPs achieve higher yield levels than LI-CMPs at high input use levels but they appear to be less efficient that LI-CMPs at low input use levels.

[9] Matrix $\boldsymbol{A}^c$ being definite positive, matrix $\boldsymbol{\Lambda}^c$ is also definite positive. In particular, a^c_{kk} and γ^c_{kk} for any input k. $f^c(\boldsymbol{x})$ yield function represents a normal technology in the sense of Rader (1968) [36] then matrix $\boldsymbol{\Lambda}^c$ is non-negative matrix (*i.e.*, $\gamma^c_{km} \geq 0$ for any pair of inputs, with $\gamma^c_{kk} \geq 0$ for any input) and all inputs are gross complements, with $\frac{\delta}{\delta w_m}x^c_k(p,\boldsymbol{w}) \leq 0$ for any input pair and $\frac{\delta}{\delta w_k}x^c_k(p,\boldsymbol{w}) \leq 0$ for any input.

[10] Our main results hold under alternative assumptions regarding matrices $\boldsymbol{A}^c$ and $\boldsymbol{\Lambda}^c$. HY techniques and nitrogen fertilization tending to increase weed and disease pressures suggest that input uses are less responsive to economic incentives with HY-CMPs than with LI-CMPs. Under the assumption stating that the HY and LI technologies are normal in Rader's sense, this would imply that conditions $\gamma^h_{km} \leq \gamma^l_{km}$ hold for $(k,m) \in K\times K$. These conditions, together with condition $\boldsymbol{b}^h \geq \boldsymbol{b}^l$, would imply that $\boldsymbol{x}^h(p,\boldsymbol{w}) \geq \boldsymbol{x}^l(p,\boldsymbol{w})$.

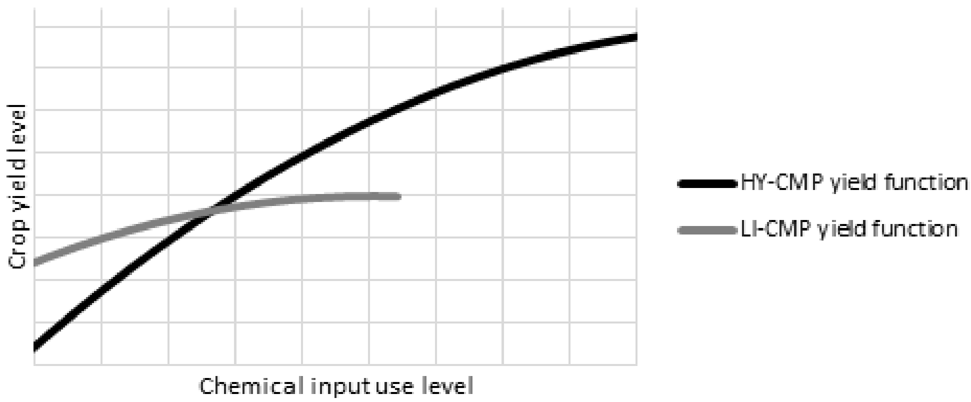


Fig. 3: Yield functions and CMPs

Comparing the optimal supply and cost functions of the LI- and HY-CMPs is straightforward since it is easily shown that:

$$y^h(p, \boldsymbol{w}) - y^l(p, \boldsymbol{w}) = b^h - b^l > 0 \tag{5}$$

and:

$$\boldsymbol{x}^h(p, \boldsymbol{w}) - \boldsymbol{x}^l(p, \boldsymbol{w}) = \boldsymbol{d}^h - \boldsymbol{d}^l > \boldsymbol{0}. \tag{6}$$

2.3 CMP adoption decision, input uses and economic incentives

Profit difference $\Delta\Lambda(p, \boldsymbol{w}) = \Lambda^h(p, \boldsymbol{w}) - \Lambda^l(p, \boldsymbol{w}) = p(b^h - b^l) - \boldsymbol{w}'(\boldsymbol{d}^h - \boldsymbol{d}^l)$ increases in p and decreases in w, implying that the relative profitability of the considered CMPs depends on the input to crop price ratios. For instance, LI-CMPs are more profitable than their HY counterparts as long as crop price p doesn't exceed the CMP switch price level given by:

$$\pi(\boldsymbol{w}) = \boldsymbol{w}'(\boldsymbol{d}^h - \boldsymbol{d}^l)(b^h - b^l)^{-1} \tag{7}$$

Results presented above can be used for illustrating the role of CMP choices on the response of input uses to economic incentives. Figure 2 is used to illustrate the following analysis. We assume, for simplicity, that farmers choose their CMP by comparing profit levels $\Lambda^h(p, \boldsymbol{w})$ and $\Lambda^l(p, \boldsymbol{w})$.[11] Let assume that p exceeds $\pi(\boldsymbol{w})$ and consider a progressive price decrease. As longs as p exceeds $\pi(\boldsymbol{w})$ farmers choose HY-CMPs that are more profitable at such price levels. Small decreases in p don't imply CMP changes and lead

[11] Of course, other factors than CMP gross returns are likely to impact farmers' decision to adopt or not to adopt LI-CMPs, including human capital, labor constraints, attitude toward risk and uncertainty, learning processes and attitude toward the environment. Yet, per technology profit comparisons lie at the root of any economic analysis of technology adoption decision.

to small decreases in chemical input uses (and in yield levels), as determined by input demand functions $x_k^h(p,\boldsymbol{w})$ (yield function $y^h(p,\boldsymbol{w})$). Yet, these decreases in crop price imply significant decreases in obtained profits and, importantly, reduce the gap between $\Lambda^h(p,\boldsymbol{w})$ and $\Lambda^l(p,\boldsymbol{w})$. At $p=\pi(\boldsymbol{w})$, farmers are indifferent between HY-CMPs and LI-CMPs. They switch to LI-CMPs when p decreases below $\pi(\boldsymbol{w})$, at least if they expect crop price to remain below for a sufficient period of time. This CMP switch leads to a drop in chemical input uses (as well as in yield levels), from $x_k^h(p,\boldsymbol{w})$ to $x_k^l(p,\boldsymbol{w})$ for input k (from $y^h(p,\boldsymbol{w})$ to $y^l(p,\boldsymbol{w})$). As p continues to decrease chemical input uses decrease at the rate described by functions $x_k^l(p,\boldsymbol{w})$ (function $y^l(p,\boldsymbol{w})$).

Interestingly, the effects of crop prices on the relative profitability of LI versus HY-CMPs points out an interesting property of LI-CMPs for risk averse farmers regarding wheat price volatility. Returns to LI-CMPs are less variable than those to HY-CMPs because LI-CMPs improve returns when prices are low and deteriorate returns when prices are high. Moreover, switching from HY-CMPs to LI-CMPs decreases the downside risk of returns to wheat production. Returns to LI-CMPs are higher than those to HY-CMPs when returns to wheat are low for both CMPs, that is to say when wheat price is low as far as price risk is considered.

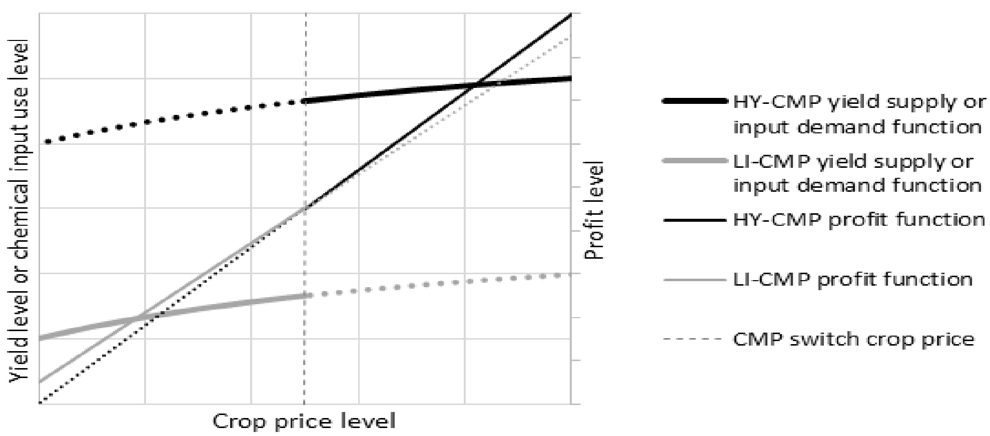


Fig. 4: Crop price, yield supply and input demand functions, and CMP choices

Let now consider the effects of input prices. Assuming that HY-CMPs are more profitable than their LI counterparts at prices $(p,\boldsymbol{w})$, taxing chemical inputs at rate α renders LI-CMPs more profitable when α exceeds the CMP switching tax rate given by:

$$\bar{\alpha}(p,\boldsymbol{w}) = p = \pi(\boldsymbol{w})^{-1} - 1 \tag{8}$$

Indeed, an increase in the price of any chemical input tend to improve the relative profitably LI-CMPs. The difference in the profit functions $\Delta\Lambda(p,\boldsymbol{w})$ is decreasing in $\boldsymbol{w}$ with

$\frac{\delta}{\delta w_k}\Delta\Lambda(p,\boldsymbol{w}) = d_k^l - d_k^h < 0$, while the CMP switch price, $\pi(\boldsymbol{w})$, is increasing in $\boldsymbol{w}$, with $\frac{\delta}{\delta w_k}\pi(\boldsymbol{w}) = (d_k^h - d_k^l)(b^h - b^l)^{-1} > 0$.

The analytical results presented in this section may explain why farmers' chemical input appear to be lowly responsive to economic incentive in the short run. Temporary changes in prices don't induce change in CMPs. They only trigger input use adjustments holding CMP choice constant. These are necessary limited as the properties of a CMP are only guaranteed in the input use range characterizing this CMP. For instance, significantly decreasing fungicide uses when using HY-CMPs is likely to lead to frequent and severe yield losses as these CMPs trigger disease risks. Similar, significantly increasing nitrogen fertilizer uses while implementing early sowing with dense seed densities, which are techniques characterizing LI-CMPs, is likely to be useless. Furthermore, this may also trigger disease risks and, as a result, enhance yield losses.

Changes in economic conditions which are expected to persistent by farmers may either induce CMP switches or comfort farmers in their current choices. Significant changes in chemical input uses occur with CMP switches.[12] Considering these features of farmers' input demands and yield supplies is important for analyzing the long run effects of economic incentives as well as modelling farmers' crop supply and input use choices. In particular, in short panel microdata farmers may use various CMPs. But, only few of them may change their CMP during the short period spanned by the data, excepted when considering easy-to-adopt and easy-to-promote techniques (e.g., new seed cultivars) or drastic and foreseeable changes in the economic context. In this case, accounting for CMPs in farmers' production choices consists in addressing the modelling and estimation issues raised by the presence of heterogeneous technologies in the considered data. Considering data covering longer periods raises other issues. Babcock (2015) [3] considered the case of time series of aggregate data. Here, we consider series of short panel micro datasets (or long rotating panel microdatasets). We have to consider the issues raised by heterogeneous technologies as well as those due to technology changes. The aggregation process smooths out the effects of these changes in macro data. But, this technology changes require special attention in micro data. They generate discontinuities, abrupt changes at least, in input use and yield series for sampled farms. These changes may only concern some farms and my occur at different points in time.

[12] Importantly, it appears to be more difficult for farmers to switch from HY-CMPs to LI-CMPs than in the opposite direction. HY-CMPs are considered as the conventional farmers. Most farmers and extension agents are experienced with these production practices. The economic and technical performances of these CMPs are well documented. By contrast, LI-CMPs were, and are still to a some extent, new practices. Adopting them entail fixed costs, mostly due to information and knowledge acquisition processes, as well as risk taking. For instance, in the *Marne crayeuse* area, which is that of our case study, performances of LI-CMPs only began to be assessed starting in the mid 90s. Dedicated experiments have been implemented all over France since 1999. As a result, if switches from LI-CMPs to HY-CMPs can be considered as short run decisions, adoption of LI-CMPs has to be considered as a medium run investment decision.

3 Empirical application: "Low input" CMPs in *la Marne crayeuse* area, 1998-2014

This section presents statistical and micro-econometric approaches aimed to infer (i) use of heterogeneous CMPs by farmers and (ii) the main characteristics of these CMPs from farm accountancy panel data. It also presents empirical results based on a rich datasets considering winter wheat production in Northeastern France. First, we present the dataset used for the empirical application. Second, we present statistical and micro-econometric approaches that we use. Third, we present our estimation results.

3.1 Data

We use an unbalanced panel dataset that considers, from 1998 to 2014, input uses and yields of winter wheat for a sample of farmers located in *la Marne crayeuse*. In our application, this dataset is used for devising 14 balanced panel datasets covering 4 years each. These (overlapping) subpanel datasets contain 342 farmers on average. These data have been extracted from cost accounting data provided by the main accounting agency dealing with farming operations in the considered area. Our application primarily makes use of recorded winter wheat yields, received wheat prices and fertilizer and pesticide expenditures devoted to wheat production. Fertilizer and pesticide price indices were obtained from the French department of Agriculture.

La Champagne crayeuse area is of special interest here for two main reasons. First, it is among the first areas in which LI-CMPs were tested with on-farm experiments (Meynard and Loyce, 1997 [23]). Moreover, these experiments were supported by local extension services and down-stream industries. At that time, local cooperatives were considering to produce ethanol from wheat. LI-CMPs were of special interest in this context. These CMPs are expected to be economically optimal at low wheat prices and ethanol production only requires low quality wheat, regarding protein content in particular.

Second, *la Marne crayeuse* is a small area characterized by homogenous agro-climatic and economic conditions.[13] This is of special interest when considering inference on CMPs from yield and input use levels. Implementation details of a given CMP strongly depends on where this implementation takes place. For example, farmers operating under favorable agro-climatic conditions and using HY-CMPs tend to target higher yield levels, and thus to use more chemical inputs, than farmers operating under less favorable agro-climatic conditions but also using HY-CMPs. As a result, the effects of exogenous factors such soil fertility and climate on yields and chemical inputs can be confounded with those of CMP choices. Also, farmers located in a small homogenous have access to the same extension services, input supply and outlets.

[13] It is one of the most productive arable crop area in France. Plots are flat and deep chalky soils are easy to plough and have very good water holding capacity. Its continental climate keeps disease pressure at relatively low levels.

Given the economic context prevailing from 1998 to 2014 and the promotion process of LI-CMPs during this period, we expect a small share of farmers using LI-CMPs in the late 90s and an increase in this share until 2006. The effects on the adoption of LI-CMPs of the combination of high wheat, fuel and fertilizer prices since 2007 is difficult to predict, notwithstanding the effects of these price volatility. If the wheat price effects dominate those of the input prices, it is possible that farmers having adopted LI-CMPs switched back to HY-CMPs.

Of course, economists tend to focus on economic drivers when analyzing farmers' technological choices. Following the seminal work of Griliches (1957) [17], adoption and non-adoption choices at a given point in time are often explained by heterogeneous returns to adoption (see, *e.g.*, Suri, 2011 [40]; Michler *et al*, 2018 [33]). Given that the farmer population considered in our study face homogenous agro-climatic and economic conditions, these heterogeneity effects are unlikely to play a major role in explaining why farmers adopted or did not adopt LI-CMPs.[14]

Recent analyses tend to highlight effects or learning – by doing or by others – processes (*e.g.*, Foster and Rosenzweig, 2010 [15]; Sadoulet, 2016 [5]), thereby pointing out the role of farmers' human capital and networks in explaining heterogeneous adoption of innovations. Howley (2015) [18], as well as sociological analyses, points out non-financial drivers of farmers' production choices, including attitude toward the environment, family concerns health concerns and taste for agronomy. These potential drivers of farmers' CMP adoption choices are unobserved. But, they may explain why farmers adopt new CMPs while others don't, despite their facing similar agro-climatic and economic conditions.

3.2 Clustering, classification and "latent CMP modelling"

Uncovering CMP choices and characteristics from farm yield and input use data consists in finding clusters of farms with similar yield and input use levels within clusters/classes and differing yield and input use levels across clusters/classes. In what follows, term y_{it} denotes the wheat yield of farmer i in year t and $\boldsymbol{x}_{it}$ denotes the corresponding vector of input use quantities.

Robust farm clustering requires considering sequences of yield and input use levels. This can easily be achieved by considering (selecting) balanced panel datasets. Considering yield supply and input demand models featuring suitable control for time varying effects (*e.g.*, weather and price effects) is necessary when using unbalanced panel datasets. In this study, we adopt the simplest approach. We construct a sequence of 14 consecutive balanced panel datasets covering 4 years each, subpanel 1998-2001 to subpanel 2011-2014, and use each of them for uncovering CMP choices and characteristics. These consecutive datatsets partially overlap as most sampled farms are observed for more than 4 years. This allows us to obtain robust clustering results per 4 year periods and to investigate the consistency of these results over time.

[14] We are gathering soil and weather information for confirming this assumption.

In what follows, term $S_{(h)} = \{t_{(h)}, ..., t_{(h)} + 3\}$ defines the set of years considered in subpanel h, with $h \in \mathcal{H} = \{1, 2, ..., 14\}$, given that term $t_{(h)}$ denotes the first year of this subpanel. If farmer i belongs to subpanel h then we observe the sequence of yield and chemical input use levels defined by $s_{i,(h)} = (s_{it} : t \in S_{(h)})$. Vector $s_{i,(h)} = (y_{it}, \boldsymbol{x}_{it})$ describes the wheat yield level and the chemical input uses of farmer i in year t.

3.3 Ascending hierarchical classification, k-means clustering and latent CMP modelling

We consider three approaches for uncovering CMP choices and characteristics from farm yield and input use data, two purely statistical methods – ascending hierarchical classification (AHC) and k-means clustering (k-MC) – and an statistical modelling approach – latent class modelling.

Clustering methods. Clustering analysis is defined as the "art of finding groups in data" (Kaufman and Rousseeuw, 2009 [19]). In our case, groups of farms are determined based on similarities and dissimilarities of yield and input use sequences. AHC and k-MC are among the main popular clustering methods.[15] The rest on the same principles. (i) They seek to group farms based on their yield and input use records such that within group variability is minimal while between group variability is maximal. (ii) They rely on standard distance functions for assessing observation similarities and dissimilarities, that is to say for assessing the distance between vectors $s_{i,(h)}$ and $s_{j,(h)}$ when farmers i and j both belong to subpanel h. AHC and k-MC differ in their algorithmic approaches. k-MC is used for obtaining allocations of farmers to a given number of groups. It reallocates farmers to groups until an optimal allocation is reached. AHC progressively forms farmers' groups, from a set of singletons to a group containing the entire farm sample.

As AHC and k-MC both yield two-cluster results consistent with a partition of each of our subpanel into a LI farmer group and a HY farmer group, we decided to focus on this two-cluster case.[16]

Latent class modelling. Although they can also be used for clustering observations, latent class modelling approaches rely on different statistical background. Identifying CMP uses and characteristics based on latent class modelling consists in specifying and estimating statistical models of farmers' yield supply and input demand functions while allowing farmers to use CMPs to be selected from a finite set of "latent CMPs". CMPs are considered latent because they are assumed to exist but neither their use nor their

[15] Applications of AHC and k-MC can be found in the agricultural and agronomic literatures. Maseda et al (2004) [27] considered AHC to distinguish Galician farms in terms of quality of life. Bellon *et al* (2001) [4], Blazy *et al* (2009) [6] and Renaud-Gentié *et al* (2014) [37] k-MC methods for uncovering CMP uses and characteristics based on detailed records of cropping operations for samples of farmers.

[16] Statistical criteria are generally used for determining the "optimal" number of groups but interpretation-focused choices are also reasonable.

characteristics are unobserved.

Let equation:

$$s_{it} = \boldsymbol{g}(\boldsymbol{z}_{s,it}, \boldsymbol{\varepsilon}_{c,it}; \boldsymbol{\gamma}_{c,0}) \text{ for } c \in C \tag{9}$$

define the yield supply and input demand function models of farmer i in year t if this farmer uses CMP c chosen from set C. Term $\boldsymbol{z}_{s,it}$ is a vector of observed production choice drivers, typically including price ratios, weather conditions and farm characteristics. Function $\boldsymbol{g}(.)$ has a known functional form in $(\boldsymbol{z}_s, \boldsymbol{\varepsilon}_c; \boldsymbol{\gamma}_c)$. It depends on parameter vector $\boldsymbol{\gamma}_c$, which characterizes CMP c. Term $\boldsymbol{\varepsilon}_c$ is vector of error terms. Assuming that (i) $\boldsymbol{\varepsilon}_{c,it}$ and $\boldsymbol{z}_{s,it}$ are independent and (ii) the probability density function (pdf) of $\boldsymbol{\varepsilon}_{c,it}$ is known up to a parameter to be estimated, $\boldsymbol{\lambda}_{c,0}$, the pdf of s_{it} conditionally on $\boldsymbol{z}_{s,it}$ and on farmer i choosing CMP c in t can be derived. Let $\nu(s_{it}, \boldsymbol{z}_{s,it}; \boldsymbol{\theta}_{c,0})$ denote this pdf, with $\boldsymbol{\theta}_c = (\boldsymbol{\gamma}_c, \boldsymbol{\lambda}_c)$.

Model (9) is standard, excepted for the fact that we don't know the CMP used by farmer i in year t. Let variable c_{it} denote the CMP used by farmer i in year t. This latent variable is discrete. Its support is given by C, which is finite. We assume, for simplicity, that farmer i uses the same CMP for $t \in S_{(h)}$. Let variable $c_{i(h)}$ define the CMP chosen by farmer i during period $S_{(h)}$. Under the supplementary assumption stating that $\boldsymbol{\varepsilon}_{i(h)} = (\boldsymbol{\varepsilon}_{it} : t \in S_{(h)})$ and $\boldsymbol{z}_{s,i(h)} = (\boldsymbol{z}_{s,it} : t \in S_{(h)})$ are independent, the pdf of the yield level and input use sequence $\boldsymbol{s}_{i(h)}$ conditionally on $\boldsymbol{z} = s, i(h)$ and $c_{i(h)} = c$ is simply given by the pdf product $\mathit{\Pi}_{t \in S_{(h)}} \nu(\boldsymbol{s}_{it}, \boldsymbol{z}_{s,it}; \boldsymbol{\theta}_{c,0})$.

Recovering parameter $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_{c,0} : c \in C)$ from the observed data requires specifying a parametric model for $c_{i(h)}$. Let the probabilistic discrete choice model of $c_{i(h)}$ be given by probability function

$$r_c(\boldsymbol{z}_{r,i(h)}; \boldsymbol{\mu}_0) = P[c_{i(h)} = c | \boldsymbol{z}_{r,i(h)}] \text{ for } c \in C. \tag{10}$$

Term $\boldsymbol{z}_{r,i(h)}$ is a vector of CMP choice potential drivers. The modelling assumptions described above yield that the pdf of sequence $\boldsymbol{s}_{i(h)}$ conditionally on $\boldsymbol{z}_{i(h)} = (\boldsymbol{z}_{s,i(h)}, \boldsymbol{z}_{r,i(h)})$ (*i.e.*, ignoring the unobserved CMP choice of farmer i) is given by:

$$\mu(\boldsymbol{s}_{i(h)}, \boldsymbol{z}_{i(h)}; \boldsymbol{\theta}_0, \boldsymbol{\lambda}_0) = \sum_{c \in C} r_c(\boldsymbol{z}_{r,i(h)}; \boldsymbol{\mu}_0) \mathit{\Pi}_{t \in S_{(h)}} \nu(\boldsymbol{s}_{it}, \boldsymbol{z}_{s,it}; \boldsymbol{\theta}_{c,0}) \tag{11}$$

where . This pdf is a suitable basis for estimating parameter based on a Maximum Likelihood (ML) approach.

Estimates of $(\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0)$ directly deliver estimates of the main characteristics of the latent CMPs, parameters $\boldsymbol{\theta}_{c,0}$ for $c \in C$, and of the frequency use of these CMPs (conditional on $\boldsymbol{z}_{r,i(h)}$), probability functions $r_c(\boldsymbol{z}_{r,i(h)}; \boldsymbol{\mu}_0)$ for $c \in C$. Together with models (9) and (10) these estimates allow constructing simulation models aimed to investigate the main driver of farmers' production choices, including their CMP choices.

Contrary to results of statistical clustering methods, estimated latent class models don't deliver farmers' CMP choices. Yet, farmers' CMP choices can be inferred from a probabilistic viewpoint. The probability of farmer i choosing CMP c conditional on $\boldsymbol{z}_{r,i(h)}$ and on the observed production choice sequence $\boldsymbol{s}_{i(h)}$ (as well as on the conditions of these choices, $\boldsymbol{z}_{s,i(h)}$), which is given by:

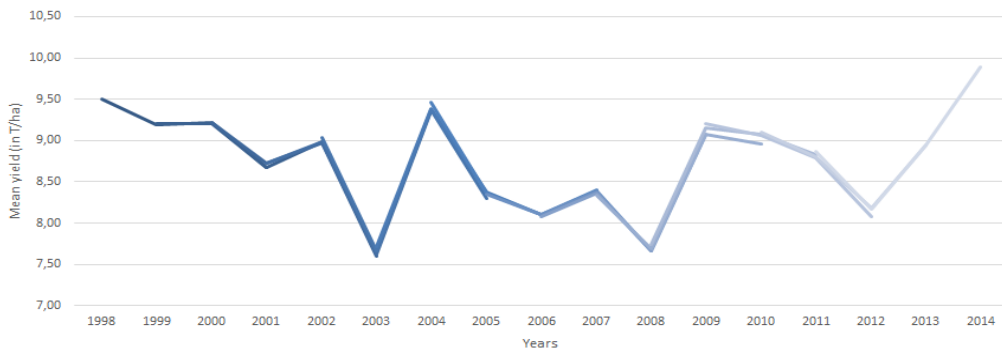
$$P[c_{i(h)} = c|\boldsymbol{z}_{i(h)}, \boldsymbol{s}_{i(h)}] = \frac{r_c(\boldsymbol{z}_{r,i(h)}; \boldsymbol{\mu}_0)\nu(\boldsymbol{s}_{it}, \boldsymbol{z}_{s,it}; \boldsymbol{\theta}_{c,0})}{\mu(\boldsymbol{s}_{i(h)}, \boldsymbol{z}_{i(h)}; \boldsymbol{\theta}_0, \boldsymbol{\lambda}_0)} \tag{12}$$

can yield precise estimates of the CMP choice of farmer i.

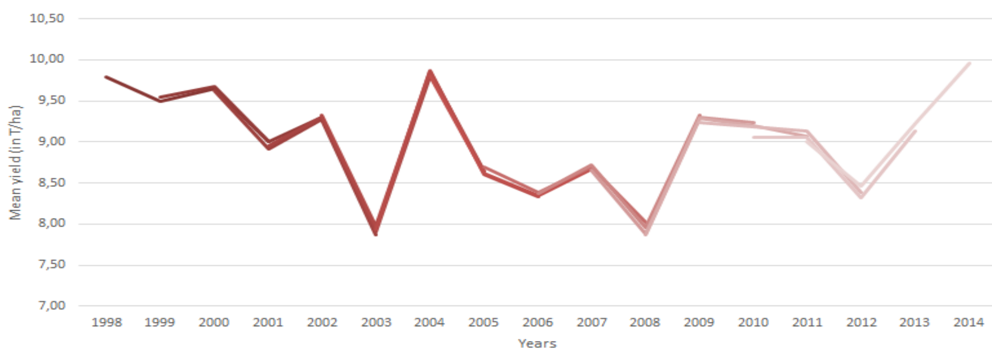
The empirical results presented in the next sub-section are obtained based on a very simple latent class model of yield level and input use sequences $\boldsymbol{s}_{i(h)}$. We assume that $\boldsymbol{s}_{it}$ conditional on $c_{it} = c$ is normally distributed with a mean vector and a variance-covariance matrix depending on h, t and c only. That is to say we assume that $\boldsymbol{s}_{it}|(c_{it} = c, h)\ \mathcal{N}(\boldsymbol{\mu}_{c,t(h),0}, \boldsymbol{\Omega}_{c,t(h),0})$ for $t \in S_{(h)}$ and $c \in C$. Similarly, we assume that $P[c_{i(h)} = c|\boldsymbol{z}_{r,it}]$ is constant, with $P[c_{i(h)} = c|\boldsymbol{z}_{r,it}] = r_{c(h),0}$. To not use any explanatory variable, year and subpanel dummy variable excepted, renders the results obtained with the "latent CMP" modelling approach comparable to those obtained with the clustering approaches.

3.4 Results

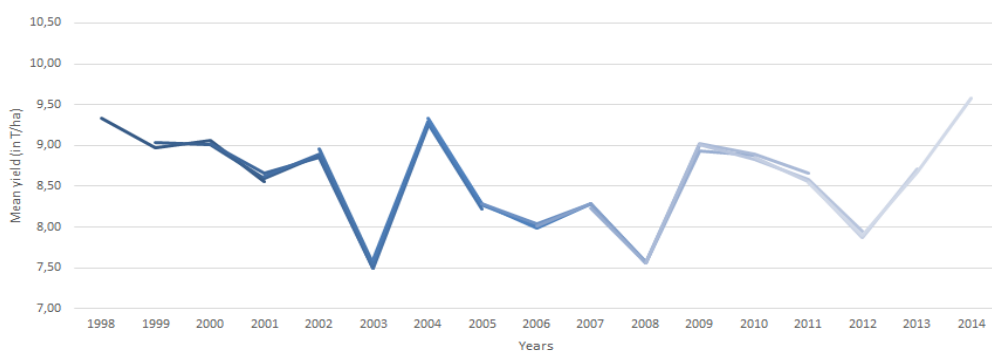
Stability of clusters over subpanels. Each analysis was performed over the 14 subpanels we extracted from the original panel data set. One concern was about the clusters characteristics through subpanels. We want that, for subpanels common years, the resulting clusters have similar characteristics in terms of yields and chemical input expenses at least when considering the same type of analyze (*e.g.* k-MC, AHC or EM). Figure 5, Figure 6 and Figure 7 allow us to verify such stability for the clusters obtained by means of cluster analyze.



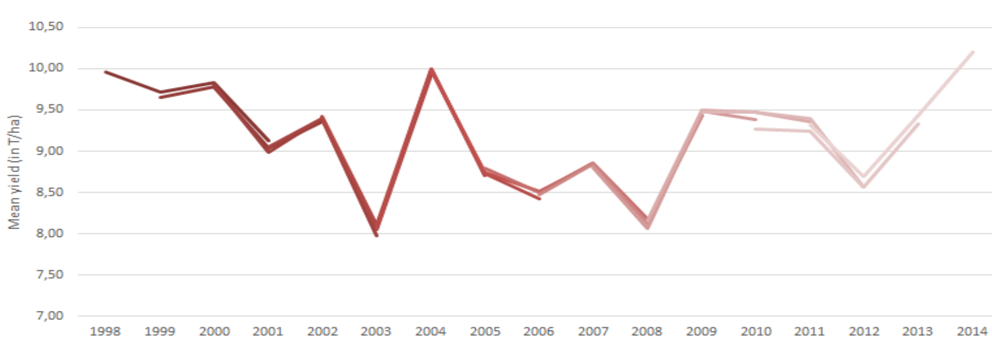
(a) k-MC - LI group



(b) k-MC - HY group

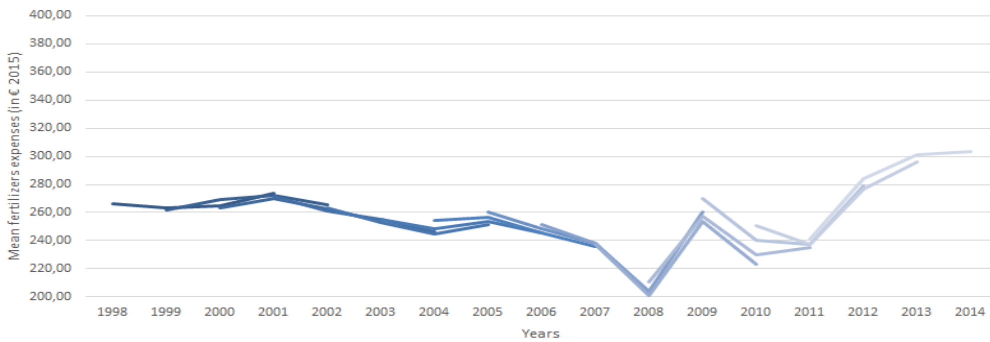


(c) AHC - LI group

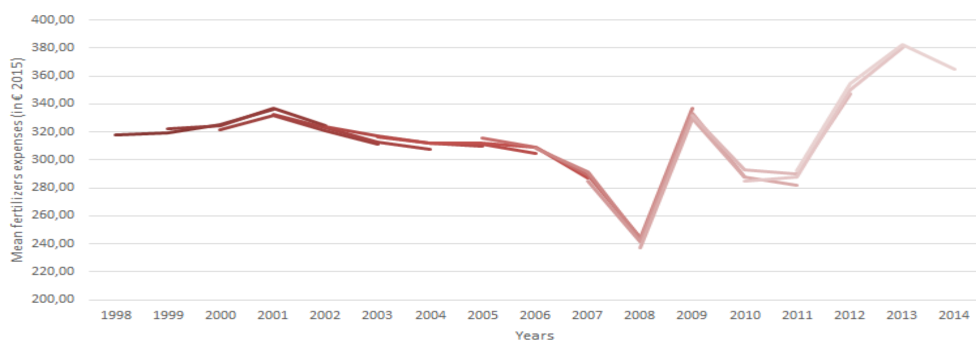


(d) AHC - HY group

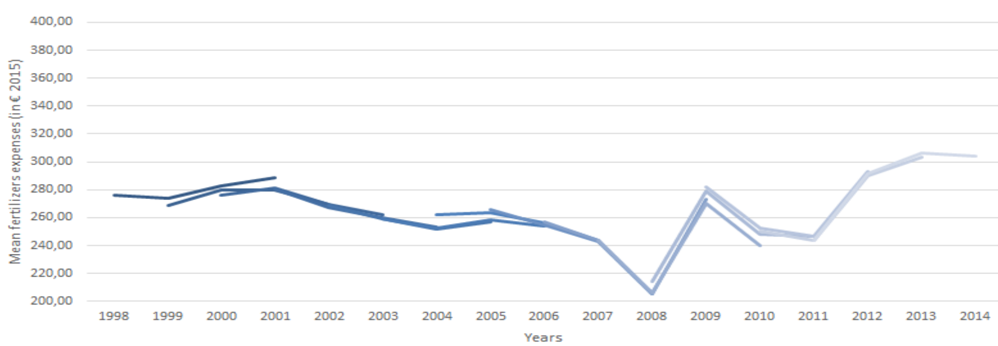
Fig. 5: Mean yield of clusters



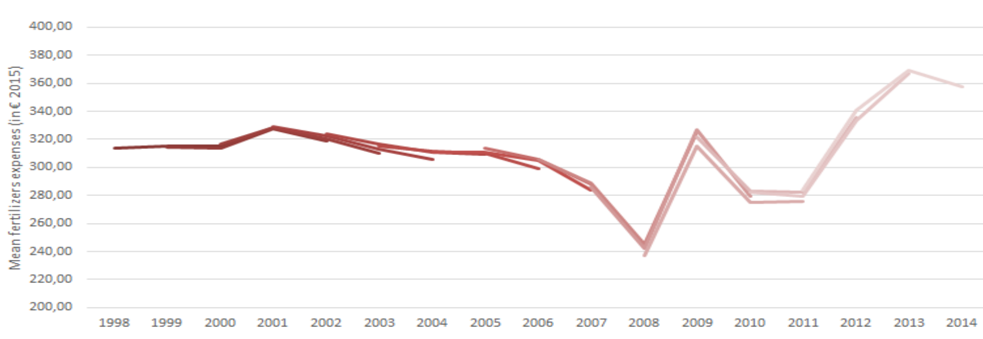
(a) k-MC - LI group



(b) k-MC - HY group

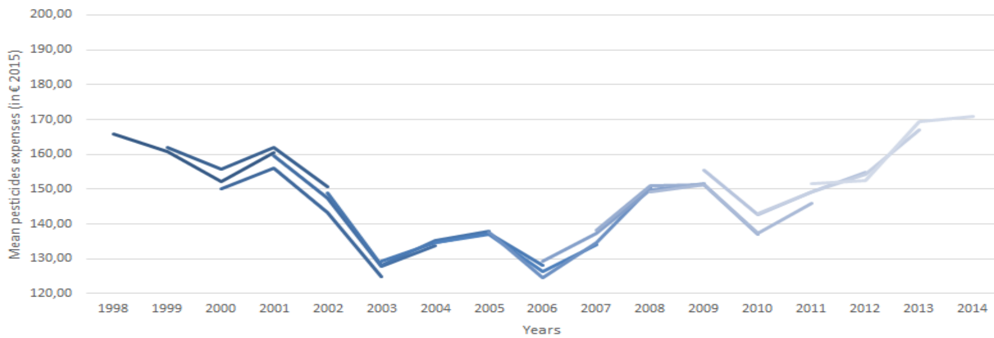


(c) AHC - LI group

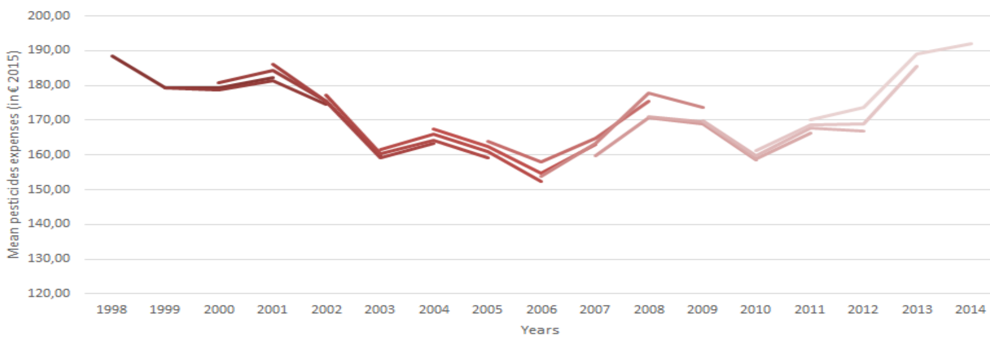


(d) AHC - HY group

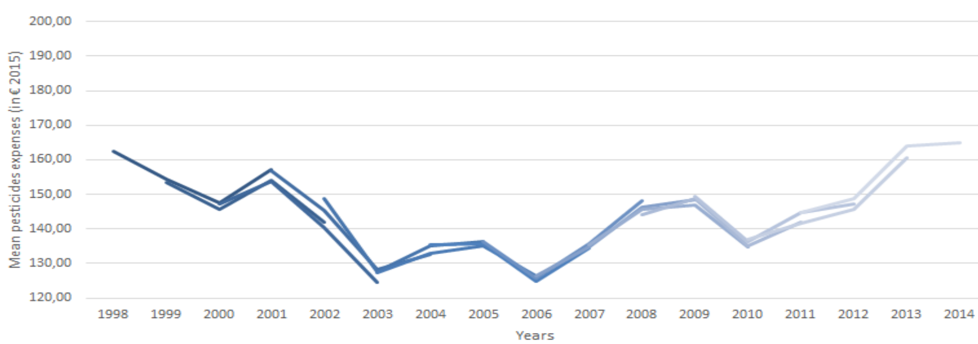
Fig. 6: Mean fertilizers expenses (in € 2015) of clusters



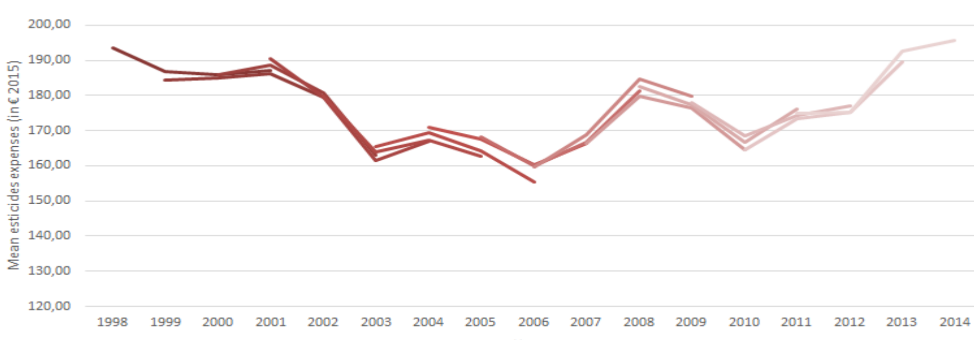
(a) k-MC - LI group



(b) k-MC - HY group



(c) AHC - LI group

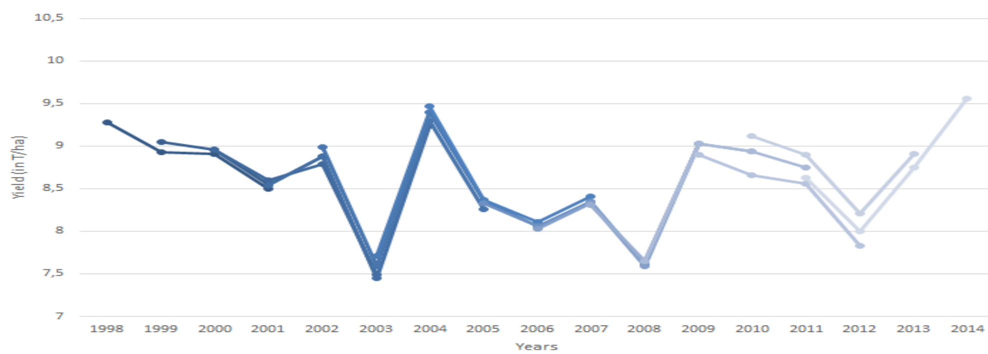


(d) AHC - HY group

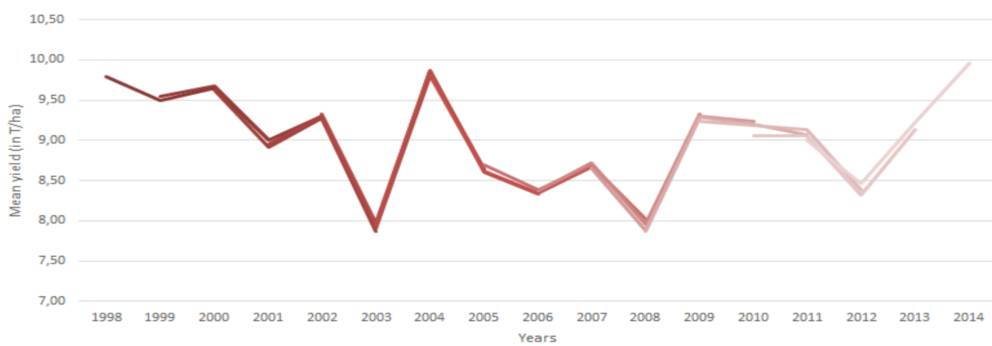
Fig. 7: Mean pesticides expenses (in € 2015) of clusters

The first observation we can make is that clusters obtained by means of cluster approaches are stable through subpanels (*i.e.* for a given year, there is no big differences in mean cluster values whatever the subpanel we consider). Another remarkable fact is that, for k-MC and AHC, the cluster profiles of yields, fertilizers and pesticides expenses are identical.

For EM results, we only represent the results obtained when using the AHC initial values (results with the k-MC initial values were similar on that point). Clusters still seem to be stable through subpanels, even if there are slight differences in mean yields and mean pesticides expenses from 2009. The observation we made previously about similar attributes profiles between k-MC and AHC results still stands when adding the EM results.

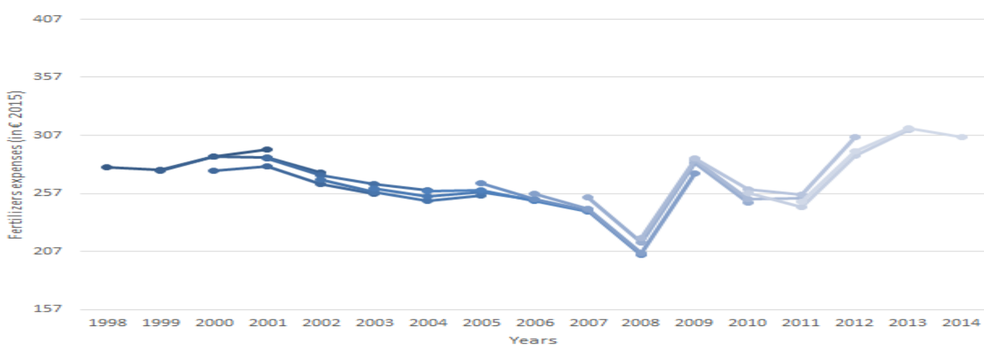


(a) EM - LI group

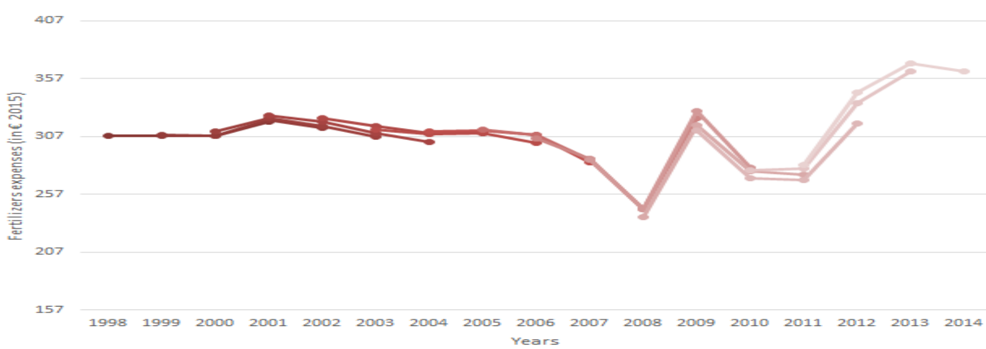


(b) EM - HY group

Fig. 8: Mean yield of clusters

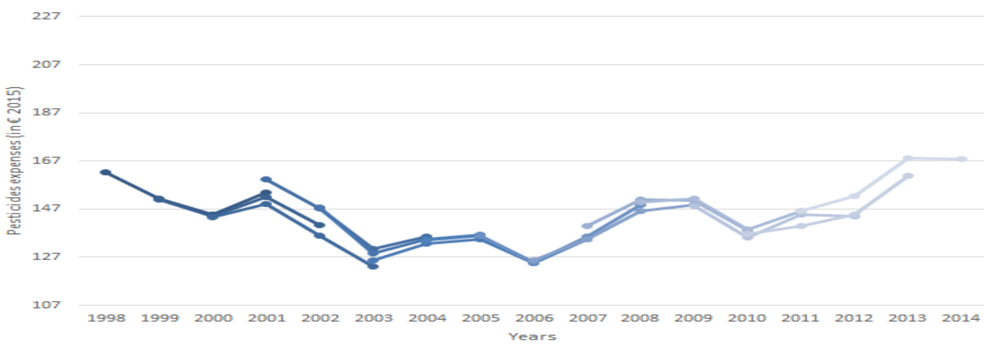


(a) EM - LI group

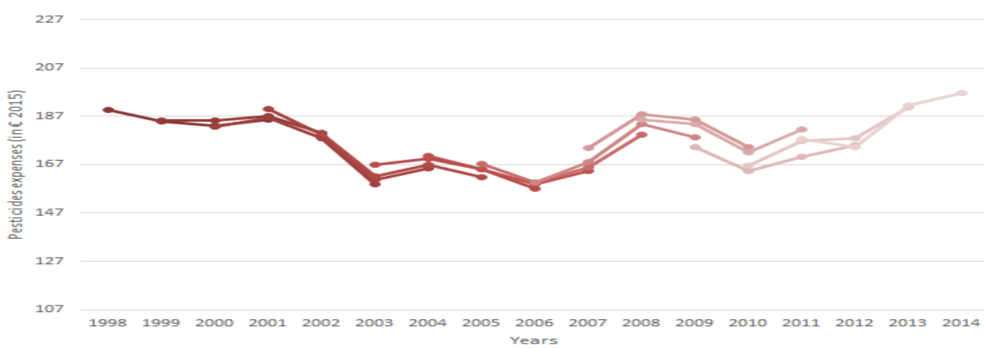


(b) EM - HY group

Fig. 9: Mean fertilizers expenses (in € 2015) of clusters



(a) EM - LI group



(b) EM - HY group

Fig. 10: Mean pesticides expenses (in € 2015) of clusters

An increasing "Low Input" group. At first, we observed the numerically composition of each cluster across time. Figure 11 shows that, whether we are considering k-MC or AHC results, the LI group is overall increasing and represents more or less 50% of the observed farms.

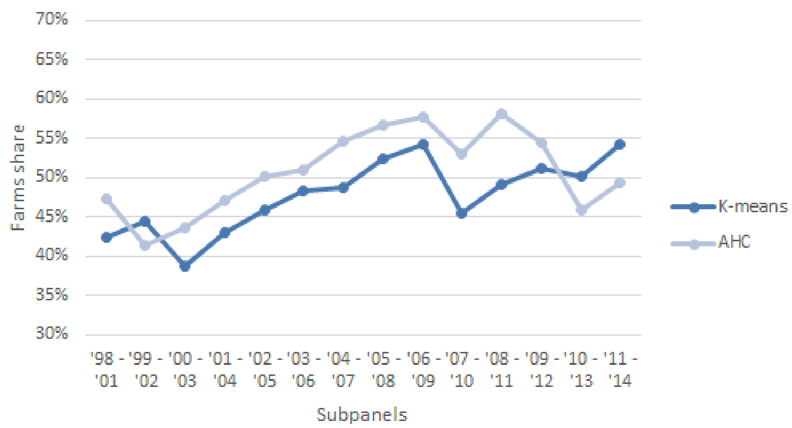


Fig. 11: Size of the LI group for AHC and k-MC approaches

Initially, we were not expecting the LI group to be almost as numerous as the HY one. The proportion of farmers adopting a LI-CMP can be understood through the on-farm experiments for LI-CMPs that were implemented in the 1990s (Larédo and Hocdé, 2014 [21]).

In this global increasing trend, we can see two inflexion points for 2007-2010 and 2010-2013 subpanels. When looking at Figure 1 and 2, we can hypothesize that those inflexion points are related to the evolution of wheat and fertilizers prices. The increase of fertilizers prices after the 2007 crises combined to the decrease of wheat prices at the same period act as an incentive for farmers to adopt "LI" CMPs. Indeed, in such price context, a CMP with a lower level of chemical inputs will minimize the farmer economic loss.

The second inflexion point observed for the 2010-2013 panel can be linked to the increase of wheat prices from 2009. The higher the wheat prices, the greater the incentive for a farmer to target high yields. Inflexion points are even more noticeable when looking at the EM results (cf. Figure 12). Until panel 2007-2010, results in terms of cluster size evolution are similar, whether considering k-MC, AHC, or EM results. We can notice that, for subpanels 2000-2003 and 2001-2004, the LI groups are slightly less numerous in the EM cases than in the clustering ones.

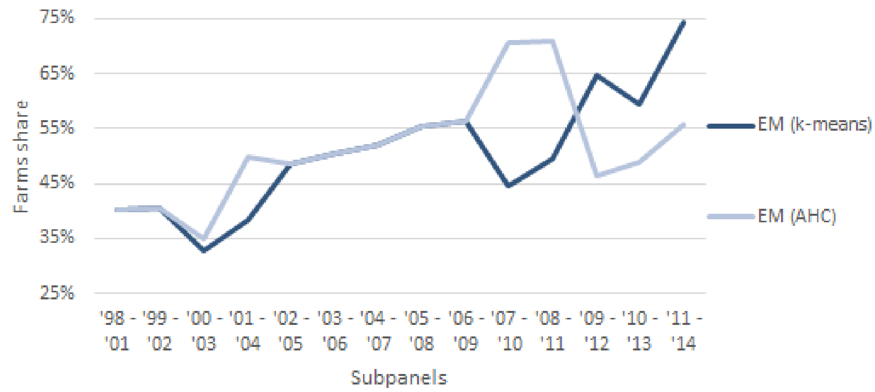


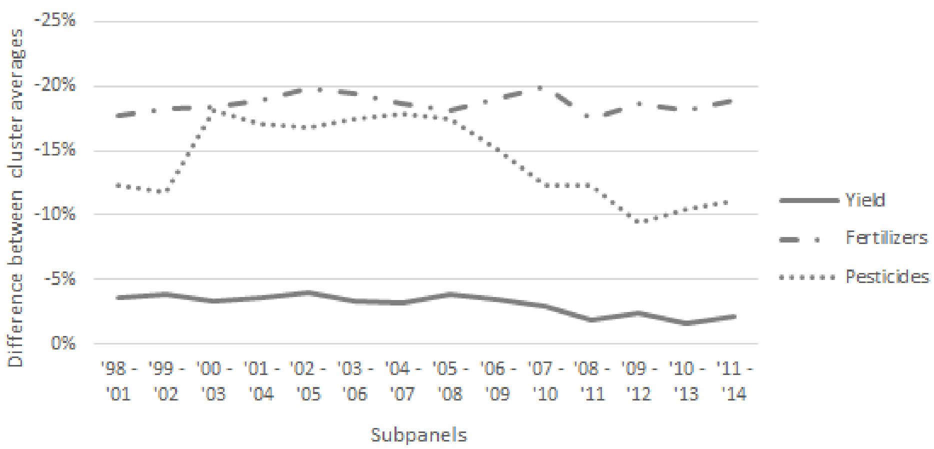
Fig. 12: Size of the LI group for EM approach

When considering the evolution of cluster size from subpanel 2007-2010, dissimilarities appear. On Figure 11, we can see that the cluster size trajectory obtained by the k-MC and the AHC are globally parallel.

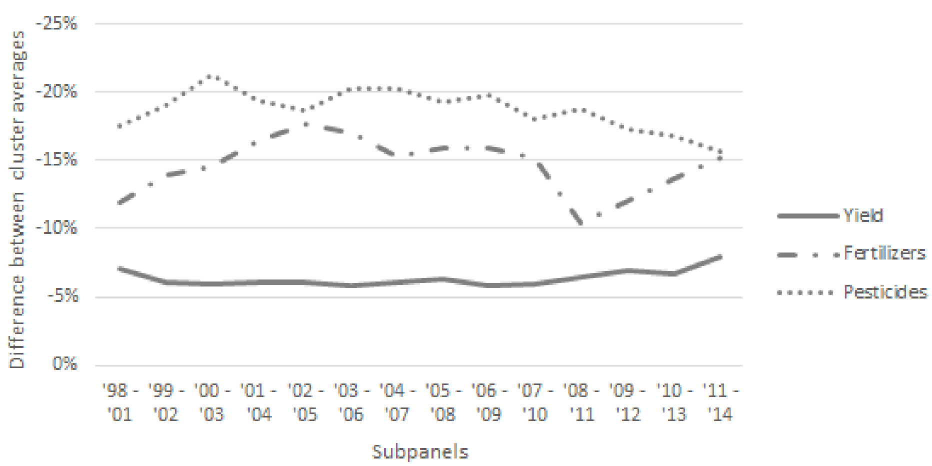
For EM, when considering either the k-MC initial values (k-MC EM) or the AHC ones (AHC EM), we can see on Figure 12 that the cluster size trajectory is opposed from subpanel 2007-2010 to subpanel 2010-2013. Cluster size of AHC EM seems to follow the trajectory of fertilizers prices. Evolution of the k-MC EM cluster size is similar to evolution of wheat prices.

Moderate differences in yields for significant input savings. The last element we have to investigate is the difference in terms of yields, fertilizers and pesticides expenses between the LI and the HY clusters. Indeed, we expect clusters to be different in terms of yields, input uses and so in their expenses.

From Figure 13, we can see that clusters are effectively associated to differences in terms of yields and chemical inputs expenses. The LI cluster is associated to a 5% yield loss compared to the HY one. This yield loss is associated to lower pesticides and fertilizers expenses: in the k-MC case, pesticides expenses are around 17% less in the LI cluster and for fertilizers expenses it is contained between 10 and 18%. In the AHC case, the LI cluster is associated to a 17% decrease in pesticides expenses and for fertilizers expenses it is contained between 10 and 16%. Even if the observed differences are a bit lower than the ones we mentioned previously (e.g. a 10% difference in yields and a 30% difference in pesticides), they are quite coherent.



(a) k-MC

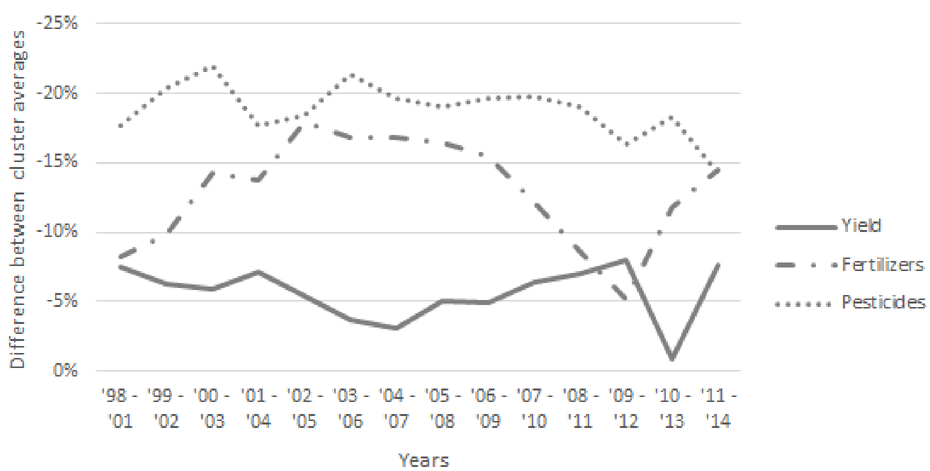


(b) AHC

Fig. 13: Relative differences between cluster means

So far, results from k-MC and AHC were very similar. However, we can see that for cluster characteristics, they are quite different. K-MC LI cluster is more about a “fertilizer saver” whereas AHC LI cluster is more about a “pesticides saver”. A priori, AHC results are more consistent with what is intended by LI CMP. If we go back to prices evolution (*cf.* Figure 1), especially the one of fertilizers, AHC results appear as more. Indeed, the drop in the yield difference that exists between the two groups corresponds to the period when fertilizers prices were high and wheat price were low. Even if they stay in the HY cluster, we expect farmers to reduce their fertilizers expenses to minimize their income loss.

Such idea is reinforced by the EM results: Figure 14 shows the same pattern as the one observed for AHC results. If we compare EM to AHC results, we can see that the differences between clusters are globally more marked when using the EM algorithm. Differences in pesticides expenses can exceed 20%.



(a) EM

Fig. 14: Relative differences between cluster means

Conclusion

An on-going work on cluster building

CMP identification results are very encouraging. Until subpanel 2007-2010, results are remarkably stable whatever the considered approach. Instability of results after 2007 encourages us to do further research in the way to improve our clustering methodology. To limit the instability propagation of few "exceptional" years to the overall cluster building we can build shorter balanced subpanels. For instance, if we consider 2007 as an "exceptional" year, in 4-years subpanels, its track will be disseminated in four subpanels. When considering 3-years subpanels, it will only affect results of three subpanels. Following such reasoning to its end leads to consider each year as a subpanel. The risk here is that cluster analyse distinguish for good and bad years. A system of weights for "exceptional years" can also be considered: to decrease the influence of outlying years, we can underweight them in the clustering process.

Globally, the unsteadiness of results after 2007 emboldens the idea that we have to account for economic effects as price changes and price volatility. Control for other non-economic factors as soil and climate conditions will ensure us that we are not differentiating for yield and input use differences due to heterogeneous conditions. Even if we expect such differences to be limited as we restricted our empirical application to farms that are situated in *la Marne crayeuse* (*i.e.* same type of soil plus similar weather conditions as farms are located in a limited area). Integration of soil and climate data is also important to run the empirical application on all farms from *la Marne*.

Refinement of the EM algorithm

So far, EM algorithm is run on each balanced subpanel independently. In the end, the idea is to run the EM algorithm on all years at the same time. To do so, we will have to make extra-assumptions on the farmers' CMP choice. The easiest way to use all the data in the EM algorithm is to assume that between 1998 and 2014, farmers always belong to the same CMP type. To determine the initial values of the algorithm, we can rely on previous clustering results and use the "majority rule". It requires that changes from a HY to a LI-CMP during the period are scarce.

If unlikely, this assumption can be relaxed in two ways. Instead of using the "majority rule", we could keep every cluster assignation obtained through subpanels and use weights for observations appearing multiple times for a given year as they belong to multiple subpanels. Otherwise, as in Markov chain approaches, transition matrices can be conceived. Transition matrices are matrices containing the probabilities to change from one state to another or to stay in the same state (*i.e.* a CMP in our case). Transition matrices can either be constant over time *i.e.* we defined transition probabilities from one CMP type to another and assume that they are time-invariant. However, in a context where incentives to reduce pesticides and to use more sustainable agricultural practices are stronger, it might seems shaky to assume such time-invariant transition matrix. This is even more accurate considering the change in the economic context that occurs between 1998 and 2014.

Another way to improve the EM algorithm as conceived now is to refine the associated production function. Until now, the production function we used was quite basic as we were focused on the CMP identification. No covariates were included and as mentioned when discussing about the clustering improvement, there are covariates that have an impact either on chemical input uses (*e.g.* relative prices of fertilizers and pesticides compared to the crop price, non-economic factors as social and moral concerns [34]), on yield (*e.g.* chemical input uses) or on both (*e.g.* weather conditions).

Exploration of technology choice determinants

Once the latent production technologies approach improved, another avenue to explore is the determinants of farmer's CMP choices. The first results of this CMP revealing tentative tends to show that economic factors have an impact on cropping practices. This idea is also supported by Wilson and Tisdell [44] as they identified economic factors as potential lockers in more sustainable practices adoption. Still, other factors influence farmers' decision as suggested by Howley [18]. Understanding the determinants of CMP choices is of particular interest in a context where the pesticide use reduction is of public interest. Indeed, bring those determinants to light might help in finding effective public policy schemes to reduce pesticides use.

Albeit preliminary, our empirical results on farm accountancy data are encouraging. Indeed, statistical and micro-econometric approaches enable us to identify two CMP types in *la Marne crayeuse*: high yielding CMPs and low input CMPs. Observed yield level differences and input use differences between clusters are similar, even if a bit lower, to what is expected by agronomists. Still, clustering results are quite stable until 2007. Adding new data (*e.g.* weather and soil data) and running the algorithms on the whole panel - instead of subpanels - might allow us (*i*) to obtain more stable clusters and (*ii*) to identify what are the determinants of CMP choice.

References

1. Alvarez, A., del Corral, J.: Identifying different technologies using a latent class model: extensive versus intensive dairy farms. European Review of Agricultural Economics **37**(2), 231–250 (2010)
2. Aubertot, J.N., Barbier, J.M., Carpentier, A., Gril, J.N., Guichard, L., Lucas, P., Savary, S., VOLTZ, M.: Pesticides, agriculture et environnement. Réduire l'utilisation des pesticides et en limiter les impacts environnementaux. Expertise scientifique collective Inra-Cemagref (décembre 2005). Expertises Collectives (2007)
3. Babcock, B.A.: Using Cumulative Prospect Theory to Explain Anomalous Crop Insurance Coverage Choice. American Journal of Agricultural Economics **97**(5), 1371–1384 (2015)
4. Bellon, S., Lescourret, F., Calmet, J.P.: Characterisation of apple orchard management systems in a french mediterranean vulnerable zone. Agronomie **21**(3), 203–213 (2001)
5. Bernard, T., de Janvry, A., Mbaye, S.: E., sadoulet, 2016. product market reforms and technology adoption by senegalese onion producers. Tech. rep., CUDARE Working Paper, UC Berkeley (http://escholarship. org/uc/item/9wj41042)
6. Blazy, J.M., Ozier-LAfontaine, H., Doré, T., Thomas, A., Wery, J.: A methodological framework that accounts for farm diversity in the prototyping of crop management systems. application to banana-based systems in guadeloupe. Agricultural Systems **101**, 30–41 (2009)
7. Böcker, T.G., Finger, R.: A meta-analysis on the elasticity of demand for pesticides. Journal of Agricultural Economics **68**(2), 518–533 (2017)
8. Bouchard, C., Bernicot, M.H., Felix, I., Guérin, O., Loyce, C., Omon, B., Rolland, B.: Associer des itinéraires techniques de niveau d'intrants variés à des variétés rustiques de blé tendre: évaluation économique, environnementale et énergétique. Le Courrier de l'Environnement de l'INRA **55**(55), 53–77 (2008)
9. Carpentier, A., Letort, E.: Accounting for heterogeneity in multicrop micro-econometric models: implications for variable input demand modeling. American journal of agricultural economics **94**(1), 209–224 (2011)
10. Carpentier, A., Letort, E.: Multicrop production models with multinomial logit acreage shares. Environmental and Resource Economics **59**(4), 537–559 (2014)
11. Féménia, F., Letort, E., et al.: How to significantly reduce pesticide use: An empirical evaluation of the impacts of pesticide taxation associated with a change in cropping practice. Ecological Economics **125**(C), 27–37 (2016)
12. Fernandez-Cornejo, J.: The microeconomic impact of ipm adoption: theory and application. Agricultural and Resource Economics Review **25**(2), 149–160 (1996)
13. Fernandez-Cornejo, J., Hendricks, C., Mishra, A.: Technology adoption and off-farm household income: the case of herbicide-tolerant soybeans. Journal of Agricultural and Applied Economics **37**(3), 549–563 (2005)
14. Finger, R., Möhring, N., Dalhaus, T., Böcker, T.: Revisiting pesticide taxation schemes. Ecological Economics **134**, 263–266 (2017)
15. Foster, A.D., Rosenzweig, M.R.: Microeconomics of technology adoption. Annu. Rev. Econ. **2**(1), 395–424 (2010)
16. Greene, W.: Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of econometrics **126**(2), 269–303 (2005)
17. Griliches, Z.: Hybrid corn: An exploration in the economics of technological change. Econometrica, Journal of the Econometric Society pp. 501–522 (1957)

18. Howley, P., Buckley, C., Donoghue, C.O., Ryan, M.: Explaining the economic 'irrationality' of farmers' land use behaviour: The role of productivist attitudes and non-pecuniary benefits. Ecological Economics **109**, 186–193 (2015)
19. Kaufman, L., Rousseeuw, P.J.: Finding groups in data: an introduction to cluster analysis, vol. 344. John Wiley & Sons (2009)
20. Khanna, M.: Sequential adoption of site-specific technologies and its implications for nitrogen productivity: A double selectivity model. American journal of agricultural economics **83**(1), 35–51 (2001)
21. Larédo, P., Hocdé, A.: Variétés rustiques et itinéraires techniques économes en intrants. Tech. rep., ASIRPA (2014)
22. Lin, B.B.: Resilience in agriculture through crop diversification: adaptive management for environmental change. BioScience **61**(3), 183–193 (2011)
23. Loyce, C., Meynard, J.: Low input wheat management techniques are more efficient in ethanol production. Industrial Crops and Products **6**(3-4), 271–283 (1997)
24. Loyce, C., Meynard, J.M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., Bernicot, M., Bonnefoy, M., Charrier, X., Debote, B., et al.: Growing winter wheat cultivars under different management intensities in france: A multicriteria assessment based on economic, energetic and environmental indicators. Field crops research **125**, 167–178 (2012)
25. Loyce, C., Meynard, J.M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., Bernicot, M., Bonnefoy, M., Charrier, X., Debote, B., et al.: Interaction between cultivar and crop management effects on winter wheat diseases, lodging, and yield. Crop protection **27**(7), 1131–1142 (2008)
26. Mahé, L.P., Rainelli, P.: Impact des pratiques et des politiques agricoles sur l'environnement. Cahiers d'Economie et de Sociologie Rurales (CESR) **4**(905-2016-70149), 9 (1987)
27. Maseda, F., Diaz, F., Alvarez, C.: Family dairy farms in galicia (n. w. spain): Classification by some family and farm factors relevant to quality of life. Biosystems Engineering **87**(4), 509–521 (2004)
28. Matson, P.A., Parton, W.J., Power, A., Swift, M.: Agricultural intensification and ecosystem properties. Science **277**(5325), 504–509 (1997)
29. McLachlan, G., Krishnan, T.: The EM algorithm and extensions, vol. 382. John Wiley & Sons (2007)
30. McLachlan, G.J., Lee, S.X., Rathnayake, S.I.: Finite mixture models. Annual Review of Statistics and Its Application (0) (2000)
31. Meynard, J.M.: Construction d'itinéraires techniques pour la conduite du blé d'hiver. Ph.D. thesis (1985)
32. Meynard, J.: Pesticides et itinéraires techniques. Phytosanitaires protection des plantes biopesticides. Paris: INRA pp. 85–100 (1991)
33. Michler, J., Tjernstrom, E., Verkaart, S., Mausch, K.: Money matters: The role of yields and profits in agricultural technology adoption (2018)
34. Mzoughi, N.: Farmers adoption of integrated crop protection and organic farming: Do moral and social concerns matter? Ecological Economics **70**(8), 1536–1545 (2011)
35. Orea, L., Kumbhakar, S.C.: Efficiency measurement using a latent class stochastic frontier model. Empirical economics **29**(1), 169–183 (2004)
36. Rader, T.: Normally, factor inputs are never gross substitutes. Journal of Political Economy **76**(1), 38–43 (1968)

37. Renaud-Gentié, C., Burgos, S., Benoît, M.: Choosing the most representative technical management routes within diverse management practices: Application to vineyards in the loire valley for environmental and quality assessment. European Journal of Agronomy (2014)
38. Rolland, B., Bouchard, C., Loyce, C., Meynard, J.M., Guyomard, H., Lonne, P., Doussinault, G.: Des itinéraires techniques à bas niveaux d'intrants pour des variétés rustiques de blé tendre: une alternative pour concilier économie et environnement. Le courrier de l'environnement de l'INRA **49**(49), 47–62 (2003)
39. Skevas, T., Lansink, A.O., Stefanou, S.: Designing the emerging eu pesticide policy: A literature review. NJAS-Wageningen Journal of Life Sciences **64**, 95–103 (2013)
40. Suri, T.: Selection and comparative advantage in technology adoption. Econometrica **79**(1), 159–209 (2011)
41. Teklewold, H., Kassie, M., Shiferaw, B., Köhlin, G.: Cropping system diversification, conservation tillage and modern seed adoption in ethiopia: Impacts on household income, agrochemical use and demand for labor. Ecological Economics **93**, 85–93 (2013)
42. Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S.: Agricultural sustainability and intensive production practices. Nature **418**(6898), 671 (2002)
43. Vanloqueren, G., Baret, P.: How agricultural research systems shape a technological regime but locks out agroecological innovations. Elsevier. Research Policy **38**, 971–983 (2009)
44. Wilson, C., Tisdell, C.: Why farmers continue to use pesticides despite environmental, health and sustainability costs. Ecological economics **39**(3), 449–462 (2001)