

Performance of innovative cropping systems diversified with oilseeds and protein crops: identification and resolution of methodological issues, using the Syppre experimental network as a case study[☆]

Sandrine Longis^{1,2}, Stéphane Cadoux³, Anne-Laure Toupet de Cordoue², Paul Tauvel⁴, Marie Estienne², Pierre Onzon³, Françoise Lescourret⁵, Clotilde Rouillon² and Jean-Noël Aubertot^{1,*} 

¹ INRAE, University of Toulouse, UMR AGIR, Centre Occitanie-Toulouse, Auzeville Tolosane, France

² Arvalis Institut du végétal, Pôle SYSTèmes de culture Multiperformants, Boigneville, France

³ Terres Inovia, Thiverval-Grignon, France

⁴ Institut Technique de la Betterave, Paris, France

⁵ INRAE, Domaine Saint Paul, site Agroparc, Avignon, France

Received 30 April 2022 – Accepted 1 September 2023

Abstract – Agroecological transition requires that innovative and diversified cropping systems be developed. Conducting system experiments is an approach well-suited to the analysis of performance of cropping systems when subjected to soil, weather and biotic stresses. Conducting system experiments nevertheless gives rise to methodological challenges. Using the Syppre network of experiments, consisting of five sites in France, we present an original case study that provides valuable methodological and agronomic lessons on system experiments. The innovative cropping systems tested there are based on crop diversification (including oilseeds and protein crops), as well as flexible tillage, technical innovations and optimized crop management. From a methodological standpoint, we show that (i) mixed models are adapted to a range of experimental questions and constraints; (ii) multifactorial analysis enables the characterization of relationships between performance indicators; (iii) a multisite experimental network is an efficient approach not only for answering agronomic questions, but also for addressing methodological issues. From an agronomic standpoint, we showed that reconciling multiple indicators of performance is still challenging. Overall, innovative and diversified systems improved the performance of input utilization and environmental impacts, but with lower productivity and profitability. Introducing legume crops is a promising strategy because this contributes significantly to reductions in mineral N fertilizer use, energy consumption and greenhouse gas emissions, without major trade-offs against other performance indicators. Finally, we showed that the nature of the production situation had a major influence on the performance profile. This led us to be cautious in making overall analyses especially with regard to general conclusions.

Keywords: Crop diversification / cropping system / experimental design / multivariate analyses / mixed models

Résumé – Multiperformance de systèmes de culture innovants diversifiés à base de cultures oléagineuses et protéagineuses : identification et résolution de difficultés méthodologiques en considérant le réseau expérimental Syppre comme étude de cas. La transition agroécologique nécessite le développement de systèmes de culture innovants et diversifiés. L'expérimentation système est une approche qui permet l'analyse des performances des systèmes de culture sous l'influence du sol, du climat et des stress biotiques. Cependant, les expérimentations système soulèvent encore des défis méthodologiques. Nous présentons une étude originale qui a permis de tirer des leçons tant méthodologiques qu'agronomiques sur les expérimentations système, à partir de l'analyse du réseau expérimental Syppre qui comporte cinq sites. Les systèmes de culture innovants testés reposent sur des successions de cultures diversifiées, avec introduction d'oléagineux et de protéagineux, ainsi que sur un travail du sol flexible, des innovations

[☆] Contribution to the Topical Issue: "Innovative Cropping Systems / Systèmes innovants de culture".

*Corresponding author: jean-noel.aubertot@inrae.fr

techniques et des itinéraires techniques optimisés. D'un point de vue méthodologique, nous avons montré que (i) les modèles mixtes sont adaptés à une gamme de questions et de contraintes expérimentales ; (ii) l'analyse multivariée permet de caractériser les relations entre les indicateurs de performance ; (iii) un réseau expérimental multisite est un dispositif efficace pour répondre à des questions agronomiques, et pour traiter différentes questions méthodologiques. D'un point de vue agronomique, nous avons montré que l'atteinte de la multi-performance reste un défi. Globalement, les systèmes innovants et diversifiés ont amélioré les performances en matière d'utilisation des intrants et d'impacts environnementaux, mais avec une productivité et une rentabilité moindre. L'introduction de légumineuses est une voie prometteuse en raison de leur contribution significative à la réduction de l'utilisation d'engrais azotés, de la consommation d'énergie et des émissions de gaz à effet de serre, sans altérer de façon majeure les autres performances. Enfin, nous avons montré que la situation de production avait une influence majeure sur le profil de performances obtenu. Ceci invite à la prudence, notamment en ce qui concerne la généralisation de conclusions d'analyses globales.

Mots-clés : Diversification des cultures / système de culture / plan d'expérience / analyses multivariées / modèles mixtes

Highlight

- Experimental design theory and multivariate statistics are useful to system experiments.
- Diversified cropping systems involved trade-off between performances, in the short term.
- Introducing legume crops improved N-related and overall performances.
- Benefits from crop diversification depended on production situations.

Introduction

Agriculture faces multiple challenges towards achieving sustainable development. It must ensure an adequate supply of healthy food (FAO, 2019) and feed (Godfray *et al.*, 2010; Foley *et al.*, 2011), while ensuring that farmers receive a satisfactory income. It must also limit its impacts on the environment (Tilman *et al.*, 2011; Silva *et al.*, 2019), biodiversity (Hallmann *et al.*, 2017; Stanton *et al.*, 2018; Brühl and Zaller, 2019) and health (Schwarzenbach *et al.*, 2010). In addition, climate change (abiotic stresses such as droughts and heatwaves, along with the development of new biotic stresses; IPCC, 2019), current global energy crisis, fertilizer scarcity and market price variability require adaptive management. Research and development activities do consider these challenges (Schiere *et al.*, 1999; Martin *et al.*, 2013). However, minor adjustments of current farming systems will not be sufficient to address the multiple challenges outlined above. The agroecological transition is a major process, one that calls for an in-depth review of management systems, in particular for strengthening ecosystem services (Zhang *et al.*, 2007; Lescourret *et al.*, 2015; Duru *et al.*, 2015; Lechenet *et al.*, 2017). Several levers can be used to enhance ecosystem services in the course of crop production. One of these is crop diversification across time and space (Gaba *et al.*, 2014), for example by introducing legumes into the crop sequence (Drinkwater *et al.*, 1998; Plaza-Bonilla *et al.*, 2017) through intercropping (Duchene *et al.*, 2017) or maximizing soil cover (Elhakeem *et al.*, 2021) with cover crops (Plaza-Bonilla *et al.*, 2017). Reducing tillage and pesticide use is another one (Vasilachi *et al.*, 2020; Cros *et al.*, 2021). However, despite many

experimental efforts, there is still a lack of knowledge about the specific effectiveness of particular combinations of levers (Duru, 2013; Rosa-Schleich *et al.*, 2019) in different production situations, a concept that includes notably soil, climate, and landscape characteristics (see the definition proposed by Aubertot and Robin, 2013, adapted from Breman and de Wit, 1983).

Examining multiple indicators of performance of management systems therefore necessitates the use of system approaches. As early as 1974, Michel Sebillotte wrote about agroecosystems: “the study of such ensembles is practically impossible with the classic experimental approach [...] since the range of combinations of factors that come into play is wide. These transformations of soil and micro-climate by crops and associated cropping practices make it necessary to study cropping systems” (Sebillotte, 1974).

Since the 1960s, system approach has been used in agricultural sciences to assess the feasibility and performance of complete management systems (Brossier *et al.*, 2012), sometimes also with the objective of better understanding the interactions within agroecosystems (Doré *et al.*, 2006). System experiments involve a wide range of action levers chosen for their expected direct or indirect effects on agroecosystem performance. They are not aimed at assessing the impact of each individual lever, even though it is sometimes possible to assess the effectiveness of individual operations (*e.g.*, mechanical weeding by comparing weed density before and after the operation).

However, although a method for conceptualizing the agroecosystem is available (Lamanda *et al.*, 2012) and the design of cropping and/or livestock systems is now well documented (Boiffin *et al.*, 2001; Meynard *et al.*, 2012; Martin *et al.*, 2013), there is little information in the literature on the design of experimental plans for system experiments. System experimenters encounter methodological problems (Schillinger 2011; Bianconi *et al.*, 2013) that are not well formalized and which can hamper experiments. Bianconi *et al.* (2013) explained that the failure to take certain elements into account when developing protocols has the effect of limiting the use of statistical methods. As a consequence, it impedes a rigorous response to the questions the experiments address. They identified the following weaknesses in particular: i) areas dedicated to system experiments too limited; ii) sample sizes too small; iii) experiment durations too long; iv) lack of a control. We propose that the value of system experiments could be improved by clearly formalizing the experimental questions in advance of

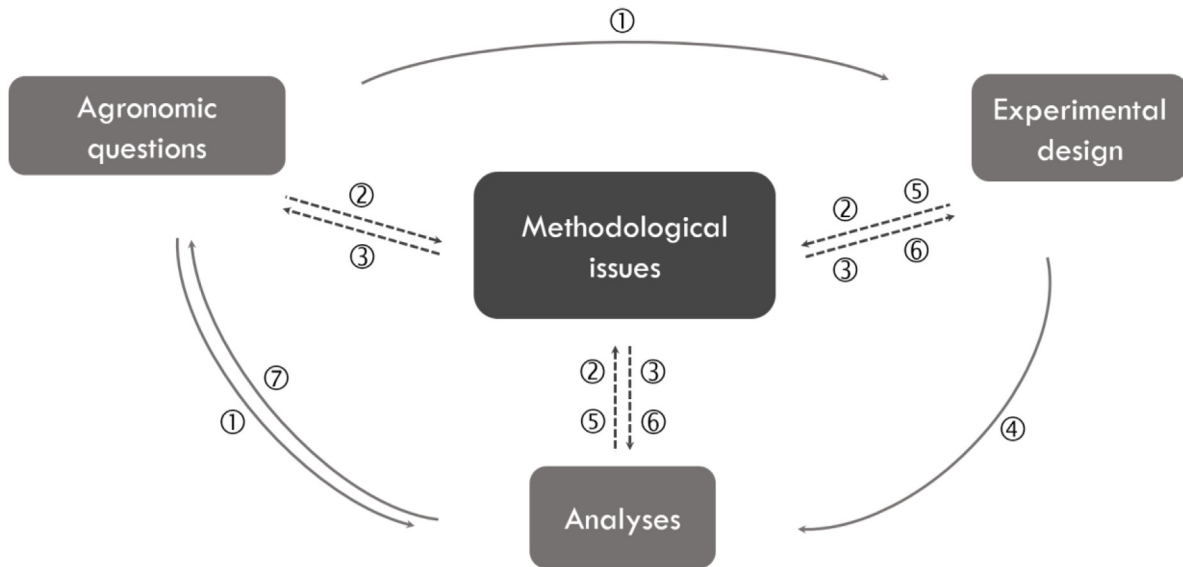


Fig. 1. Simplified conceptual diagram of the main issues of system experiments in agronomy. Experimental design must be conducted in accordance not only with agronomic questions, but also with relevant analyses (1). At each step, the identification of methodological issues (2) and their resolution (3) enables reliable analyses (4) that can lead to new agronomic questions (7). When carrying out data analyses, new methodological issues may arise (5) and require adaptations (6).

engaging in experimental design. More generally, the links between methodological issues related to system experiments and agronomical questions (Fig. 1) are so strong, that they should be dealt with holistically.

In order to address both the agronomic and the methodological issues, this paper uses a case study based on a dataset obtained from a multisite network system experiment implemented within the French project named Syppre (Tauvel *et al.*, 2019; Viguier *et al.*, 2021). Syppre deals with annual arable crops and seeks to test innovative cropping system strategies that diversify the crop through introducing oilseed and protein crops into the rotation.

In this paper, we present the main methodological issues raised by the design of the Syppre experimental network and the agronomic questions it addresses. With regard to the latter, we have selected three priority questions: (i) With regard to productivity, profitability, input use and environmental impacts, do innovative cropping systems perform better than the control systems in the production situations under consideration? (ii) Are there conflicts between performance indicators? (iii) Do diversification and the cultivation of oilseed and protein crops result in improved cropping system performance?

We then propose methods for answering these questions and then apply them to the Syppre dataset in order to analyze and discuss the agronomic questions.

Materials and methods

Description of the Syppre experimental network

Syppre, a collaborative ongoing project initiated in 2014 (Toqué *et al.*, 2015; De Cordoue *et al.*, 2016, 2018; Dubois *et al.*, 2019), consists of an experimental network of diversified cropping systems (Cadoux *et al.*, 2019; Tauvel *et al.*, 2019) implemented in five locations that are

representative of France's major regions for arable crop production: Picardie (PIC), Champagne (CHA), Berry (BER), Lauragais (LAU) and Béarn (BEA; Fig. 2). These different situations enable an examination of a large range of soils, climates, crops and value chains. In each site (Fig. 2), two cropping systems were implemented: one innovative containing at least one oilseed or protein crop in the rotation; the other a control. The objective of this experiment was for the innovative systems to: (i) verify their technical feasibility; (ii) assess their performance, (iii) fine-tune them. Overall, the experiment aimed at enhancing knowledge on agroecological cropping systems.

Cropping system design was carried out by local expert working groups. Regarding the innovative cropping systems, their design was carried out according to the *de novo* design method (Meynard *et al.*, 2012), based on a prototyping process (Vereijken, 1997) conducted in workshops. The innovative cropping systems were designed to meet, both general multipurpose objectives, common to the five production situations (Tab. 1), and address local issues, specific to each production situation (Tab. 2). *Ex ante* assessments of performance were made along the lines of those described by de Cordoue *et al.* (2016). An iterative design process was then carried out until an innovative cropping system had been selected which, *a priori*, could meet the objectives.

The control cropping system was defined by regional groups drawing together their collective expertise and the results of regional surveys on crop sequences and cropping practices. These control systems were designed to be representative of the dominant crop rotation in the region under consideration, with optimized cropping practices to ensure high technical and economic performance. Following Debaeke *et al.* (2009), once a consistent set of decision rules to trigger technical operations had been formulated, the two cropping systems were implemented in each of the five experiments. The innovative cropping systems were widely

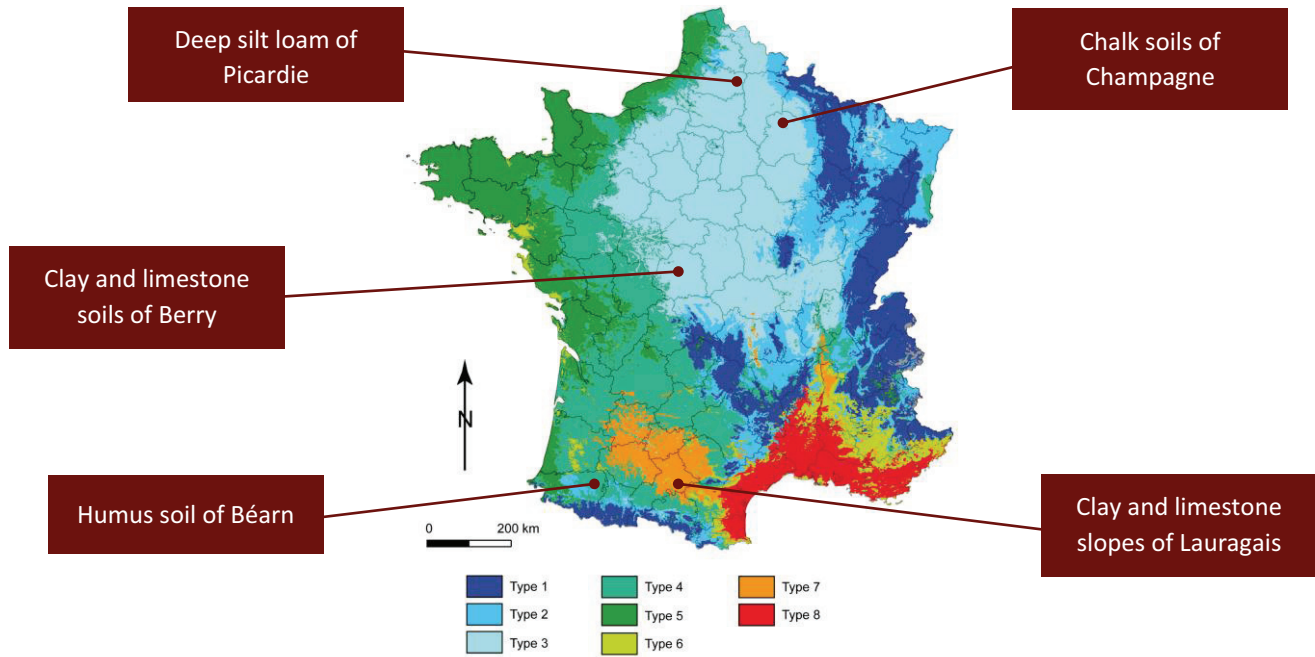


Fig. 2. Geographical location of the Syppre experimental platforms on a climate map (Joly *et al.*, 2010) with soil type according to geographical regions: Béarn, Berry, Champagne, Lauragais, Picardie. The types correspond to the following climates. Type 1: mountain climate; Type 2: semi-continental climate and the mountain margins climate; Type 3: Modified oceanic climate of the Central and Northern Plains; Type 4: Semi-oceanic climate; Type 5: Oceanic climate; Type 6: Semi-mediterranean climate; Type 7: Southwestern Basin climate; Type 8: Mediterranean climate.

Table 1. Overall objectives common to the innovative cropping systems of the five platforms of the Syppre experimental network: issues, indicators, and target. *The carbon stock will be used only for the final evaluation of systems.

Issues	Indicators (unit)	Targeted result
Productivity	Gross energy production ($\text{MJ} \cdot \text{ha}^{-1}$)	$\geq$ control system
	Energy efficiency (unitless)	$\geq$ control system
Profitability	Direct margin ($\text{€} \cdot \text{ha}^{-1}$)	$\geq$ control system
	Earnings Before Interests Taxes Depreciation Amortization (EBITDA) per Human Work Unit (HWU; $\text{€} \cdot \text{HWU}^{-1}$)	$\geq$ control system
Input use and environmental impacts	Treatment Frequency Index (TFI; unitless)	-50% compared to regional reference
	Amount of mineral nitrogen fertilizer ($\text{kg} \cdot \text{ha}^{-1}$)	-20% compared to control
	Primary energy use ($\text{MJ} \cdot \text{ha}^{-1}$)	-20% compared to control
	Greenhouse Gas emissions (GHG; $\text{teqCO}_2 \cdot \text{ha}^{-1}$)	-20% compared to control
	Carbon stock ($\text{t} \cdot \text{ha}^{-1}$)*	$\geq$ control system

diversified as compared to the control systems, the number of crops being between 75 and 250% greater depending on the platform (Tab. 2).

The experimental set-up was based on blocks repeated two or three times depending on the platform concerned in order to ensure the robustness of the results. In each block, all crops in the rotation of both systems were present each year and were randomized to avoid bias from crop/year interactions (Fig. 3).

Each unit plot (one crop of one system in a given year; Fig. 4) was between 12 and 24 meters wide and between 45 and 75 meters long. This was in order for farming equipment to be used and thus be cultivated under conditions as close as possible to those typically encountered. Each year, cropping practices were applied according to sets of technical decision rules and were recorded. The experiments commenced in 2016 and are designed to have a duration of minimum ten years. As

Table 2. Description of production situations and main characteristics of the systems implemented in each platform of the Syppre experimental network. B: spring barley; CP: canning pea; DW: durum wheat; FB: faba bean; GM: grain maize; LE: lentil; PO: potato; S: Sorghum; BW: Buckwheat; SB: sugar beet; SF: sunflower; SoB: Soybean; SP: spring pea; WB: winter barley; WP: winter pea; WR: winter rapeseed; WW: winter wheat; cc: cover crop; cce: harvested cover crop for energy purpose; lIm: legume living mulch; ”/”: multiple crops; “+”: intercrops.

Platform	Production situation	Control system	Main strategies of the innovative system	Local challenges
Picardie PIC	- Type 3 climate - Deep silt loam - Non-irrigated - Arable and industrial crops	- Conventional farming - Six-year rotation: ccSB-WW-ccPO-WW-ccCP or WR+lIm-WW (4 main crops; 1 oilseed or protein crop) - Alternating plowing and reduced tillage	- Lengthened (nine-year) and diversified rotation: ccSB-WW-ccFB-WW-ccGM-ccCP-WR+lIm-ccPO-WW (7 main crops; 1 oilseed; 2 protein crops) - Insertion of legumes as main and cover crops - Diversified cover crops - Flexible tillage with no plowing	- Exploiting high production potential - Mineral nitrogen (water quality, GHG emissions) - Soil fertility (storage of organic matter and reduced susceptibility to slaking and compaction)
Champagne CHA	- Type 3 climate - Deep Chalky soils - Non-irrigated - Arable and industrial crops	- Conventional farming - Five-year rotation: WR-WW-ccB-ccSB-WW (4 main crops; 1 oilseed; 0 protein crop) - Alternating plowing and reduced tillage	- Lengthened (ten-year) and diversified rotation: WR+lIm-WW-ccB-ccSB-WW-ccWB+WP-ccSF-WW-ccSB-SP (7 main crops; 2 oilseeds; 2 protein crops) - Insertion of legumes as main and cover crops - Diversified cover crops - Reduced tillage	- Production – including biomass - Production quality - Mineral nitrogen (GHG) - Soil fertility (organic matter, N dependency, slaking, erosion)
Berry BER	- Type 3 climate - Shallow and stony clay-limestone soil - Non-irrigated - Arable crops	- Conventional farming - Three-year rotation: WR+lIm-WW-WB (3 main crops; 1 oilseed; 0 protein crop) - Reduced tillage	- Lengthened (nine-year) and diversified rotation: ccLE-DW-WR+lIm –ccGM-ccSF-WW-WP+WW-WW-WB (8 main crops; 2 oilseeds; 2 protein crops) - Insertion of legumes as main and cover crops - Diversified cover crops - Flexible no-till or reduced tillage	- crop robustness against weather hazards and pests - Soil fertility (nutrient availability) - Weed control
Lauragais LAU	- Type 7 climate - Clay-limestone slopes - Non-irrigated - Arable crops	- Conventional farming - Two-year rotation: SF-DW (2 main crops; 1 oilseed; 0 protein crop) - Alternating plowing and reduced tillage	- Lengthened (eight-year) and diversified rotation: WR+lIm-DW+lIm-WB+lIm-ccS-WP/BW-DW-ccSF-WW (7 main crops; 2 oilseeds; 1 protein crop) - Insertion of legumes as main and cover crops - Diversified cover crops - Multiple cropping - Flexible no-till or reduced tillage	- Production and quality - Erosion and soil fertility (organic matter) - Inputs (water quality, GHG)
Béarn BEA	- Type 4 climate - Humus-bearing clays - Non-irrigated - Arable crops	- Conventional farming - GM monoculture (2 main crops; 0 oilseed; 0 protein crop) - Plowing	- Two innovative cropping systems implemented (I1 and I2): - Lengthened (2 and 3 year) and diversified rotation. I1: ccGM- WB/SoB-WW/cce (3 main crops; 0 oilseed; 0 protein crop) I2: cce/GM- ccSoB (2 main crops; 0 oilseed; 1 protein crop) - Multiple cropping	- Controlling weeds and soil pests with - Plant protection products - Species diversity

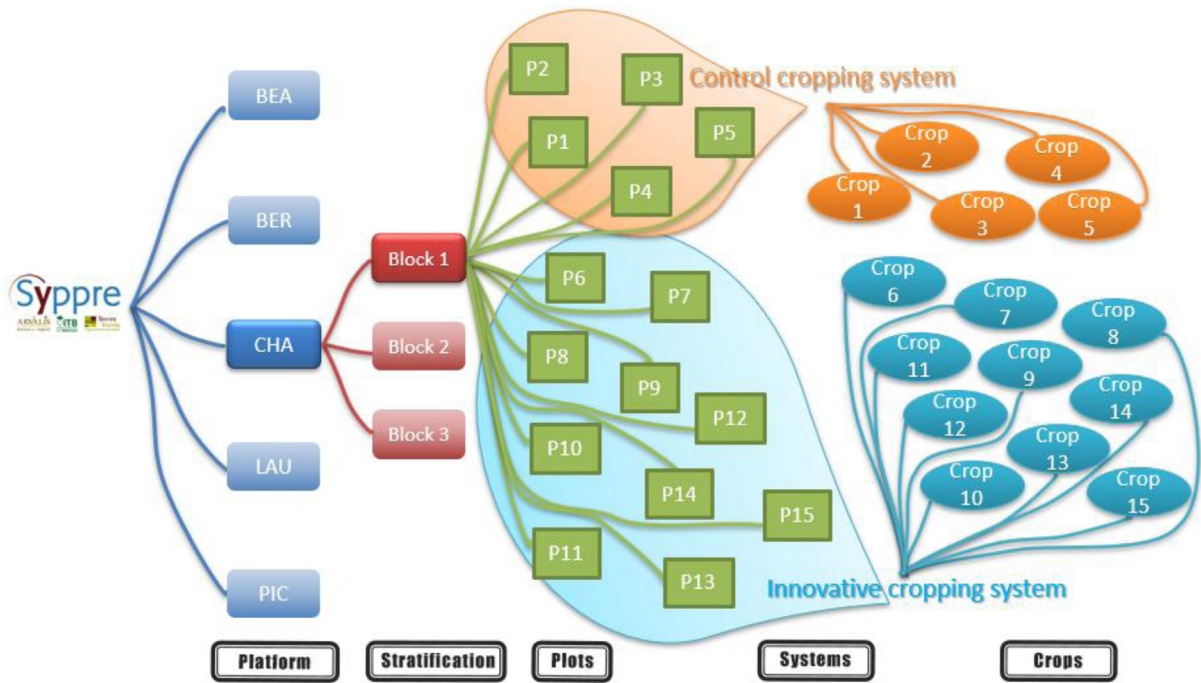


Fig. 3. Conceptual diagram of the experimental design of the Syppre's experimental network. It is a simplified representation of the components and their interactions in a given year. It is based on the models presented by (Zuur *et al.*, 2009) in order to translate the physical and technical components of the system experiment into statistical components. The block 1 structure is representative of the two other blocks (repetition of the experiment), and the CHA (Champagne) platform is representative of the four other platforms (BEA: Béarn, BER: Berry, LAU: Lauragais, PIC: Picardie). Plots are elementary units on which measurements are made. They were randomly assigned to a cropping system. An initial crop was randomly assigned to each plot at the beginning of the experiment. All crops follow each other on each plot across the years according to the crop sequence. In order to retain the diagram's clarity, the randomization of the experimental layout is not presented.

		Temporal dimension			
		Year 1	Year 2	Year 3	
Spatial dimension	Plot 1	Crop 1	Crop 2	Crop 3	Crops sequence over time in a given plot
	Plot 2	Crop 2	Crop 3	Crop 1	
	Plot 3	Crop 3	Crop 1	Crop 2	

Crops distribution in a given year

Fig. 4. Simplified representation of the spatial and temporal dimensions of cropping systems. The diagram is based on a theoretical situation with only three plots in a block and three crops in the considered rotation.

pre-crop effects have a major impact on crop performance, it was decided not to use the 2016 starting year in the analysis of the results. The four years considered were 2017, 2018, 2019 and 2020.

Performance indicators

The detailed analysis and comparison of cropping system performance was based on eight of the nine design objective indicators (Tab. 1; the carbon stock will be used only for the final evaluation of cropping systems). Two additional indicators were also considered: total working time—to add a social dimension to the analysis —, and the quantity of pesticide active ingredients to complete the pesticide use evaluation. The design objectives indicators were selected to reflect the overall performance objectives, focusing on three

main aspects: (i) productivity to meet an increasing demand for food, feed, energy and materials; (ii) profitability for farmers; (iii) low use of inputs and low environmental impacts to minimize the overall footprint of cropping systems (Tab. 2). A gross energy production indicator was used as a proxy for the overall estimate of photosynthesis, while energy efficiency was used to gauge energy production in a context of limited energy availability. Direct margins per hectare were taken as the indicator of profitability, with EBITDA (Earnings Before Interests Taxes Depreciation Amortization) used to estimate farmers' incomes. The Treatment Frequency Index (TFI), the amount of mineral nitrogen fertilizer, and the energy consumption were used to assess inputs. TFI expresses the frequency of pesticide treatments and is often considered as a proxy for a pesticide risk indicator (Kudsk *et al.*, 2018). TFI was calculated as the summation of ratios of applied rates of

each active substance divided by its standard approved rate, weighted by the proportion of treated field area. Greenhouse gas (GHG) emissions including direct and indirect emissions were chosen to consider an overall environmental impact. The soil carbon stock indicator was not selected in the analysis because it refers to a slow biogeochemical process and its evaluation after only five years did not appear relevant.

Diversification index, applied to overall rotation diversification as well as diversification through oilseeds crops and protein crops was calculated at the block level according to [Keichinger *et al.* \(2021\)](#). Implemented cropping practices and all other variables were collected at the plot level using the SYSTERRE[®] tool ([Weber *et al.*, 2019](#)). Performance indicators were calculated using the same tool (SYSTERRE[®], [Weber *et al.*, 2019](#)). The calculation of energy and GHG emission indicators was based on emission factors, taking into account direct and indirect emissions/consumptions, defined within the framework of the GESTIM project ([Gac *et al.*, 2010, 2011](#)), then a conversion into CO₂ equivalent, all according to IPCC guidelines ([IPCC, 2006](#)).

The overall dataset does not have any missing values.

Identification of methodological issues and methods to overcome them

Our methodological approach had three main steps:

- 1 Take stock of the available knowledge (objectives, protocol, experimental design)
- 2 Identify methodological issues
- 3 Propose solutions to the identified methodological issues and a method to answer the agronomic questions

The identification of methodological issues was carried out in two distinct stages based on a dialogue between analysts and experimenters, completed by methodological considerations supported by the conceptual scheme presented in [Figure 3](#). First, we listed methodological issues encountered by experimenters upstream of analysis. Second, we completed this list with other issues encountered during subsequent methodological reflections (while considering solutions, data processing).

Reaching solutions to methodological issues was based on a conventional statistical approach whereby experiments were organized in a loop of actions applied to our conceptual model ([Fig. 1](#)): “planning/implementation/analysis/interpretation” ([Dagnelie, 2012](#)). A methodological difficulty identified at one stage may find a solution at another stage of the experimental process. We chose to begin at the point where the agronomic questions were formalized because these were fundamental to the experimental design and the choice of analyses. As [Dagnelie \(2012\)](#) rightly pointed out: “Data must be consistent with the objectives and experimental questions of the considered trial. That is, they must be thought of since the design of the protocols”. Thus, we analyzed the agronomic questions in order to assess:

- 1 Whether they were aimed at extrapolation or not. Thus, whether or not statistics should be used to answer them;
- 2 If so, whether the question concerns a comparison, an evolution or a relationship between variables ([Tab. 3](#)), in order to refine the choice of statistical methods to be used. Solutions to methodological issues were then proposed according to whether they required ([Fig. 1](#)):

- (i) a modification to the agronomic question; (ii) a revision of the experimental design and/or data management; (iii) the identification of a relevant data analysis method. This choice was made according to an identification key for analytical methods made for this work ([Tab. 3](#)).

Four types of methodological issues were identified: i) data entry and presentation (between agronomic question and design, [Fig. 1](#)); ii) specificity of the agronomic and statistical vocabularies (between agronomic question and analyses, [Fig. 1](#)); iii) planning of experiments; iv) specificities of system experiments.

Since the identified methodological issues were primarily associated with agronomic questions, we chose to present them according to these questions. [Table 4](#)—simultaneously presents the list of methodological issues, as well as proposed solutions based on the method presented in [Table 3](#).

The column “actions” enables us to specify, when necessary, the method that was implemented for the analysis of the Syppre dataset to overcome the experimental issues.

One observes that the methodological questions common to all agronomic questions ([Tab. 4](#)) concern experimental design. A better knowledge of experimental design is a big step towards solving methodological issues. In our case, the actions proposed in [Table 5](#) led to the writing of the equation (1): a mixed model that answers the first agronomic question and can be applied to each indicator.

$$I_{ijkl} = \mu + (S_i + SY_{ij}) + (P_k + B_{i(k)} + SP_{ik} + SYP_{ijk}) + \epsilon_{ijkl}(1). \quad (1)$$

The first bracket contains the fixed effect factors and the second bracket contains the random effect factors. Each variable of equation (1) is defined as follows:

- I_{ijkl} : mean value of indicator I , at level i of factor S (system), at level j of factor Y (year), at level k of factor P (platform) and at level l of factor B (block)
- μ : overall mean, quantifying the influence of: i) the two fixed effects studied (system and year); ii) the controlled random effects (platform and block); iii) and their interactions allowed by the experimental design.
- S_i : measures the mean difference induced by the level i of factor S . The System effect is a fixed effect.
- Y_j : measures the mean difference induced by the level j of the factor Y . The Year effect is a fixed effect.
- SY_{ij} : measures the mean interaction difference between factor S at level i and factor Y at level j . This effect, generated by two fixed effects, is also a fixed effect.
- P_k : measures the mean deviation induced by the level k of the factor P . The Platform is considered as a random effect factor.
- $B_{i(k)}$: measures the mean deviation induced by the level l of the factor B . Block is considered as a random effect factor nested in the k levels of the random effect factor P (Platform).
- SP_{ik} : measures the mean interaction difference between level i of factor S and level k of factor P . This effect is random due to the randomness of the Platform.
- SYP_{ijk} : measures the mean interaction difference between level i of factor S , level j of factor Y and level k of factor P .

Table 3. Simplified table representing the identification key of possible analytical methods (Gomez and Gomez, 1984; Siegel and Castellan Jr, 1988; Cady, 1991; Federer, 1999; De’ath, 2002; Baayen *et al.*, 2008; Zuur *et al.*, 2000; 9; Dagnelie, 2012; Payne, 2015) based on three generic experimental questions (column 1). This table does not enable one to answer agronomic questions on a case-by-case basis, but it does provide guidance and access to the main families of data analysis whenever possible. It is read from left to right. The user chooses at each step the line that corresponds to a given experimental situation. The proposed analyses can be undertaken only after having verified that the data meet their validity conditions. A good use of this table begins with the identification of the experiment’s statistical terms: factors, modalities, measurement scale, blocks, control.

Objective	Key 1	Key 2	Main families
Compare groups or systems	1 factor	2 groups / modalities	Comparison tests of 2 samples – According to scale of measurement and size
		>2 groups / modalities	Comparison tests of more than two samples – According to scale of measurement and size
	>1 factor	No blocks	Model, multi-factor ANOVA type
		Blocks	Mixed models
Follow an evolution over time	Time factor only	No control	Graphics only (confounding effect with weather effect)
		Control	Quantitative measurements –> Longitudinal models
			Qualitative measurements –> Longitudinal models –> Proportional tests
	>1 factor	Quantitative measurements	Longitudinal models
		Qualitative measurements	Longitudinal models Proportional test
Study the relationships between variables / indicators	No causal link	2 variables	Quantitative & ordinal measurements –> Correlation tests Nominal measurements –> Contingency coefficients –> Correspondence Analysis
		>2 variables	Concordance test Multi-factorial analyses
	Causal link	2 variables	Quantitative measurements –> Regression
			Qualitative measurements –> Proportion comparisons (chi2) – –> Correspondence Analysis
			Mixed measurement scales –> Sample comparisons
		–>2 variables	Quantitative measurements –> Multiple regression –> Principal Component Analysis
			Qualitative measurements –> Discriminant analyses –> Multiple correspondence analysis
			Mixed measurement scales –> Model to specify –> Mixed Factor Analysis

Table 4. Main methodological issues common to the three agronomic questions addressed with system experiments.

Methodological issues	Solutions	Actions
Should we consider that the cropping system is applied either: - on a plot with a different sequence term (temporal dimension) - on all the plots that each receive a sequence term (temporal and spatial dimensions)	If there is only one plot per year, the studied system has only a temporal component. If the system has several crops present each year, the choice will be made according to whether or not hypotheses consider the interaction between crops and/or the overall contribution of all crops to the considered indicators of performance.	Manipulation of data so that rows in the data table correspond to the statistical units associated with agronomic questions: the plot or all the plots of the cropping system. If the format of the data frame has already been considered during the planning phase, and that the database easily allows extractions, this operation can be performed with ease.
Should we use data collected at the plot level or at the block level?	It depends on the question. To optimize the relationship between variables/indicators, the scale must be as close as possible to the underlying hypotheses of the statistical analysis.	Tables 5, 6 and 7 show that to answer the three agronomic questions, we used the two levels.
Can data from the different experimental sites in the network be gathered, and a cross-cutting analysis be conducted?	Yes, if there is a common protocol and measurements/observations. It is important not to give identical names in the platforms for elements that are in fact different. For example, a block can only be found in one platform.	The platforms become modalities of a "platform" variable in a global data table. Give individual names to the blocks and plots to make the models fit the experimental design
Are there enough replicates?	This question should be addressed during the planning phase Power tests are very difficult to implement in system experiments, given the amount of data collected and variables/indicators on which tests can be applied. Practical constraints rarely allow for acceptable number of replicates, but it is advisable to get as many as possible in order to avoid attributing to cropping system performance that are in fact random. The number of replications is evaluated by taking into account the spatial and temporal units that can be processed as blocks.	Identification of all the components of the experimental layout: experimental sites, blocks, plots. This permits the identification of the smallest statistical unit (here, the plot) and to deduce the number of units under the same experimental conditions and the total number of units involved in a given statistical model.
Can we answer agronomic questions that can arise after the planning phase?	This aspect has to be taken into account when planning the experiment. This can be achieved by ensuring that the experiment has a control and sufficient replicates, and a database system sufficiently flexible to allow any new variable to be stored. Thus, data already present could be used to answer new questions. This question is crucial for long-term experiments.	For each new question, consider whether the data collected are appropriate to answer it. This was the case for the question on the benefits of diversification with oilseeds and protein crops.

This effect is random due to the randomness of the Platform.

- ϵ_{ijkl} : measures the difference between the value of the experimental unit $ijkl$ and the value predicted by the model ϵ_{ijkl} model error, is what remains unexplained by the model. The deviations ϵ_{ijkl} are random.

This model enables us to evaluate the effects of year (which embeds confounded multiple effects such as weather and biotic stresses), the type of cropping system and their interaction, while controlling the random effect variables “experimental platform” and “block”, according to the crossings and nesting’s highlighted during the development of the conceptual diagram ([Fig. 3](#)).

These methodological results also highlight the need for flexibility in data tables or databases. The performance indicators that can be extracted from Syppre’s experiments are extremely numerous and not all are pertinent to answering the considered agronomic questions.

Statistical analyses

All the crops of each system are implemented each year in the Syppre experimental network. This allows two ways for analyzing data ([Fig. 4](#)). First, one can consider the temporal dimension of the system under study. This consists of analyzing the data obtained on a given plot during a complete

Table 5. Main methodological issues and solutions to answer the question "Do innovative cropping systems perform better than the control systems in the production situations under consideration?".

Methodological issues	Solutions	Actions
Should we use data collected at the plot level or at the block level?	In order to control, as well as possible, for possible sources of bias in the comparison of the two cropping systems, it is advisable to use the smallest unit. We choose the plot here because the difference between plots will be found in the random variation, whereas with the blocks it could be confused with the effect of crops and not necessarily be found in the random variation evaluated between blocks.	
How to compare cropping systems?	Given the complexity of cropping systems, the first step is to verify that there are no elements that could distort the comparison by confounding effects or by increasing random variation. If it is a cross-cutting analysis, the different levels of spatial scale must be controlled. A mixed model is the most suitable statistical model for this question, provided that its conditions of application are respected.	Check that items that could increase random variation cannot be controlled even if they were not included in the experimental design. Use of mixed models to control the site and block effects Lattice graphs are well-suited for visualizing all the effects that apply to the dependent variable.
How to write a model to compare systems without bias?	A good knowledge of the experimental layout to identify the relevant point.	Design of the conceptual scheme of the experiment; identification of statistical units, random effect variables, fixed effect variables, the way modalities intersect or interlock.
In the context of an experimental network, can we compare innovative systems with control systems if they do not have the same number of systems in each block?	It is necessary that the modalities of the fixed effect variables are balanced between blocks. If not, it is advised to retain only common modalities for carrying out analyses.	The Béarn platform tests three cropping systems (one control and two innovative) whereas other platforms only consider two. To overcome this agronomic issue, we considered only one of the innovative systems of this platform. The choice was made for the most promising system (I2, Tab. 2).
How to implement temporal variables in statistical models?	If one of the objectives is to study trajectories of cropping systems' effects, time is to be considered here as a fixed factor, whose effects are to be evaluated. When this is not the case, time can be considered as a random effect factor, allowing repetitions within the experiment.	One of the objectives of the Syppre experimental network is to study trajectories of the effects of cropping systems. Time was therefore considered here as a fixed factor to be evaluated. A mixed model was implemented where time was considered at the interaction 'platform x cropping system' level.
How to compare cropping systems with different rotations, possibly with different durations?	If crop diversity is not a lever and crops should not bias the comparison, the crop variable can be considered as a random variable and controlled in a mixed model. If diversification is considered as a lever, and that one system is more diversified than another, the "crop effect" should not be controlled in the model, as this would affect the comparison of systems Duration should be a controlled factor. Its impact can be estimated at any time, but with specific interpretations at the end of each cropping sequence.	Implement mixed models without the "crop variable" being controlled. Our analyses only show the effect of time during the sequences because no system has yet completed its crop rotation.
How to deal with heteroscedasticity, notably caused by heterogeneity of crops with regard to variability, for comparison of cropping systems?	It is necessary to use models whose heteroscedasticity is corrected by a mathematical transformation of the dependent variable. This is not always satisfactory.	Scaling transformation for dependent variables when necessary. Use of logarithm, square root and inverse functions, alone or in combination.

Table 6. Main methodological issues and solutions to answer the question “Are there conflicts between performance indicators?”.

Methodological issues	Solutions	Actions
Should we use data collected at the plot level or at the block level?	Plot is the preferred statistical unit because it is the unit on which the measurements/ observations were made and the most reliable for studying correlations.	
How to study discrepancies between performance indicators?	Analyze non causal relationships between many variables. Multifactorial analyses are therefore appropriate.	Correlation analysis and Principal Component Analysis (PCA) in our case, since all indicators are quantitative.
How to select variables/ indicators that will be active variables in a multifactorial analysis?	Clearly draw the distinction between result variables/indicators (dependent variables) and explanatory variables/indicators (independent variables) describing: (i) cropping practices and (ii) the production situation. Remove from the dataset all redundant and unrelated indicators used in the analysis.	
How can we easily identify discrepancies and concordances between indicators when an increase in a performance indicator can mean either an improvement or a deterioration of the considered performance depending on the specific situation?	Transform indicators whose increase indicate a lower performance by adding a minus sign. Rename the transformed indicators.	We have added a minus sign in front of indicator values whose increase reveals a poorer performance. The names of these indicators begin with “m.”

rotation. Second, one can consider the spatial dimension of the system under study. This consists of analyzing the data obtained on all the crops of the rotation in a given year. We favoured spatial dimension analyses because: (i) cropping systems are fine-tuned over the years; (ii) the time required to observe cumulative effects of cropping systems is not known *a priori*.

In order to compare innovative systems against control systems for all the sites, a cross-cutting analysis was carried out by indicator. Since the number of indicators was considerably high (53), we focused on the ten indicators presented above (see Sect. 2. Performance indicators). The statistical unit considered was the plot to ensure the most accurate comparison possible.

A principal component analysis was used to study links between indicators (additional indicators included) considered plots as statistical units. A total of 53 indicators that do not present any redundancy ($r > 0.95$) were used as active variables (Table S1).

A second principal component analysis was conducted to study performance, using blocks as statistical units. It was based on 47 selected result indicators. Moreover, a set of 126 variables/indicators (including diversification indices) were used to describe the performance classes. The complete analysis was thus performed with a set of 173 variables/indicators (Table S1). In addition, a hierarchical clustering on principal components was performed to describe links between cropping practices and performance. One can remark that the two statistical units considered (plots and blocks) led to different numbers of indicators (53 *versus* 47, respectively). This is because only those satisfying the conditions for using PCA were kept since a given variable or indicator has different number of observations/measurements depending whether plots or blocks are considered as

statistical units. In addition, some indicators were specific to the statistical unit considered and therefore also contributed to the discrepancy between the numbers of indicators associated to each type of statistical units considered (*e.g.*, yield is specific to the plot level; see Appendix 1 for the list of indicators).

Basic statistical analyses were performed using common functions of the R software (R Core Team, 2020). More specific analyses were performed using the following functions:

- `lmer{lmerTest}`, for the implementation of mixed models (extension of simple linear models containing both fixed and random effects and random effects);
- `xyplot{lattice}`, for the creation of lattice graphs associated to mixed models;
- `PCA{FactoMineR}`, for the implementation of Principal Component Analyses (unsupervised machine learning method that reduces the number of variables of a data set, while preserving as much information as possible);
- `plot.PCA{FactoMineR}`, for the display of graphs of Principal Component Analyses;
- `HPCP{FactoMineR}`, for the implementation and visualization of hierarchical clustering on principal components (unsupervised machine learning method to group data into hierarchical clusters in the form of a tree);
- `catdes{FactoMineR}`, for the description of categories, including those resulting from a clustering operation.

Statistical analyses implemented were preceded by the verification of the validity conditions in particular: models residuals homoscedasticity, verification of the PCA correlations (no redundancy, interpretation of the variables only when their $\cos^2$ –relative contribution- were greater than or equal to 50% for both dimensions studied to avoid interpretation errors).

Table 7. Main methodological issues and solutions to answer the question “Do diversification and the cultivation of oilseed and protein crops result in improved cropping system performance?”.

Methodological issues	Solutions	Actions
Should we use data collected at the plot level or at the block level?	The statistical unit is the block since the question requires consideration of crop diversity in the crop sequence.	We used a data frame presenting the mean per block in a row.
How to summarize performance indicators by block?	The method should be adapted to each experimental situation. In particular, we can use the median, which is not sensitive to outliers, the mean if outliers make sense in the summary, the sum if one does not want a value that is weighted by statistical individuals. If blocks do not have the same number of plots, sums will be affected.	The control and innovative cropping systems were not implemented in the same number of plots. The choice of crop sequence is therefore important for the overall evaluation of the systems. We chose the mean to summarize information since it takes into account the number of plots. The value of an indicator for a block-year corresponds to the mean of its values in the plots of the block.
How to describe crop diversification, oilseed and protein crop rate when there are no specific indicators?	Data extracted from the database do not always provide variables immediately relevant for answering specific questions. Need for manual input: creation of variables to describe oilseeds, protein crops and diversification (see below) from the names of cultivated species in the database.	Calculate diversification, oilseed and protein crop rate in each block each year from the plot data table.
How to describe diversification?	Use simple calculation of the ratio between the number of different crops and the number of crops in the crop sequence. Use of a diversification index (Keichinger <i>et al.</i> , 2021).	
What calculation should be used to describe the rates of oilseeds and protein crops?	A simple relationship between the number of oilseed/protein/crops and the number of crops in the crop sequence. Adaptation of a diversification index (Keichinger <i>et al.</i> , 2021) to oilseeds and protein crops based on their proportion in the crop sequence.	
What analysis should be conducted to measure the association between performance indicators, diversification and oilseed and protein crop rates?	Analyze non causal relationships between many variables. Multifactor analysis conducted on selected variables/indicators is relevant to illustrate multidimensional performance. Diversification index, oilseed and protein crop rates are descriptive variables. A hierarchical ascending classification on the factorial coordinates is thereafter recommended.	Principal Component Analysis (PCA) on a data table where a row is a block measured in a given year. PCA performed on the result indicators. Hierarchical clustering on the components of the PCA. Description of the classes with all indicators, as well as diversification indices, oilseed rate and protein crop rates.

Results

Performance of innovative cropping systems as compared to the control cropping systems

We present below the results of the mixed models when applied to the indicators identified as priorities in the Syppre project (Tab. 8). Overall, we first note that the year effect was significant for almost all the indicators (except for total working time). However, the year x system interaction was never significant (Tab. 8). This indicates that the indicators have undergone a variation over the years that was equivalent for the control and innovative systems.

In terms of productivity and profitability, the performance of the innovative systems was generally worse than those of the control systems. Gross energy production was significantly lower in the innovative systems (Tab. 8). This result was mainly explained by a lower productivity of the diversification crops included in the innovative systems and sometimes by a lower yield for the same crop, as was the case for example in Picardie for potato and sugar beet (Tab. 9). The gross energy production of the innovative system in Béarn was equivalent to that of the control (Fig. 5). This is explained by the fact that the low productivity of soybean (as compared to maize in the crop sequence of the control) was compensated by the multiple

Table 8. Results of the mixed models applied to the 10 main performance indicators. We indicated which transformation had been applied to each indicator in the model, then the p -value associated with the year effect and the year x system interaction. We provide more detailed results for the cropping system effect: Fisher's F value and p -value. The last column is an aid to interpretation. EBITDA: Earnings Before Interests Taxes Depreciation Amortization. y_i : individual value for the considered variable. y : set of values in the entire dataset for the considered variable. y_0 represents the various unit variables ($y_0 = 1$ unit of y_i) used to standardize y_i , when y_i is not dimensionless and used as an argument of a transcendental function, or added to the dimensionless numerical value 1. ICS: Innovative cropping system; CCS: Control cropping system.

Indicator (y)	Transformation used	Year p -value	Year x System p -value	p -value system	F System	Direction of the difference between ICS and CCS
Gross energy production (MJ.ha ⁻¹)	$\sqrt{y_i}$	<1e-03	0.806	<1e-03	17.1716	ICS < CCS
Energy efficiency	$\sqrt{(y_i + 1)}$	<1e-06	0.580	0.640	0.2243	ICS $\cong$ CCS
Direct margins with aids (€·ha ⁻¹)	$\log\left(\frac{y_i}{y_0} - \min\left(\frac{y}{y_0}\right) + 1\right)$	0.025	0.776	0.090	3.8534	ICS $\cong$ CCS
EBITDA per unit of human work time (€·HWU ⁻¹)	$\sqrt{\left(\frac{y_i}{y_0} - \min\left(\frac{y}{y_0}\right) + 1\right)}$	<1e-04	0.269	0.029	15.9577	ICS < CCS
Total working time (h·ha ⁻¹)	$\sqrt{\left(\frac{y_i}{y_0} - \min\left(\frac{y}{y_0}\right) + 1\right)}$	0.160	0.239	0.472	0.5327	ICS $\cong$ CCS
Treatment Frequency Index (unitless)	$\log(y_i + 1)$	<1e-02	0.776	0.101	5.4448	ICS $\cong$ CCS
Quantity of active pesticide ingredients (g·ha ⁻¹)	$\sqrt{y_i}$	<1e-02	0.994	0.820	0.0611	ICS $\cong$ CCS
Amount of mineral nitrogen fertilizer (kg·ha ⁻¹)	No transformation	0.031	0.770	<1e-06	39.8222	ICS < CCS
Total primary energy used (MJ·ha ⁻¹)	$\log\left(\frac{y_i}{y_0}\right)$	<1e-03	0.936	0.030	35.2472	ICS < CCS
Total GHG emissions (teqCO ₂ ·ha ⁻¹)	No transformation	<1e-02	0.682	<1.e-05	35.6729	ICS < CCS

cropping of biomass oat. For Béarn, the higher variance for the innovative system was explained by the difference in gross energy production from the different crops of the crop sequence (soybean *versus* biomass oat + maize), whereas the control system is based on a monoculture of maize. Similarly, the higher variances of the cropping systems in Béarn, Picardie and Champagne are certainly explained by a higher difference among crops, with crops high levels of energy production: oats + maize; sugar beet and potato; sugar beet, respectively.

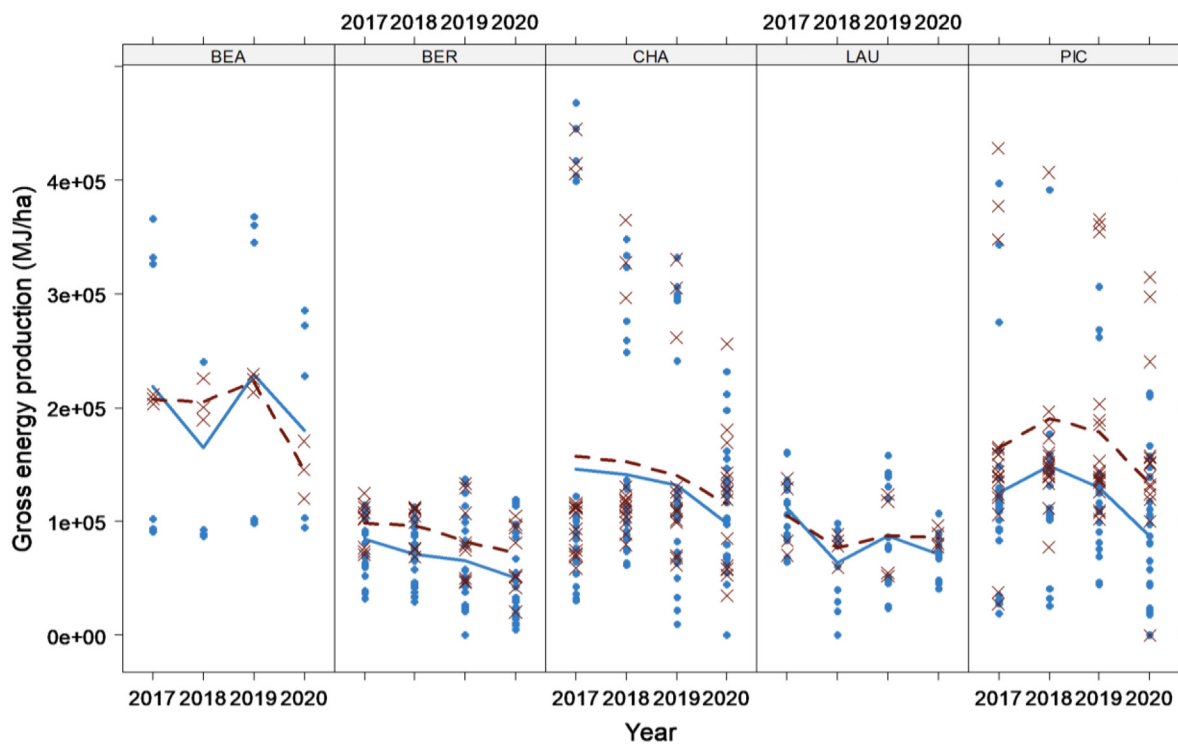
The EBITDA per unit of human labor time was also significantly lower in the innovative systems than in the control systems (Tab. 8). This was not the case for the direct margin with aid, but the low p -value does not allow us to conclude that the two systems were equivalent. The overall pattern was one of lower profitability among the innovative systems. However, graphical analysis (Fig. 6) point to differing situations across the platforms. The direct margins were clearly lower for the innovative systems in Picardie and Lauragais, slightly lower in

Berry and Champagne and equivalent, or even higher, in Béarn.

In terms of input use, technical and environmental impact indicators, the innovative cropping systems performed better overall than the control cropping systems. Treatment Frequency Index (TFI) and the quantity of pesticide active ingredients were not significantly lower in the innovative systems (Tab. 8). However, a graphical analysis reveals differences between platforms. TFI were lower for most of the innovative systems, except for Lauragais (Fig. 7). This specificity of Lauragais was explained by the difficulty of reducing the TFI of a current system based on a short rotation of crops (*i.e.*, durum wheat and sunflower), less impacted by pests than other arable crops. Regarding the quantity of pesticide active ingredients, the performance of the innovative systems was diminished compared to the TFI criterion. For this criterion, the two systems from Béarn, Berry and Champagne were equivalent and the innovative system in

Table 9. Potato and sugar beet yields in the Picardie experimental platform for the innovative system and the control, for the four years analyzed.

		Year			
		2017	2018	2019	2020
Potato	Control system (t/ha)	45.0	53.4	55.6	29.0
	Innovative system (t/ha)	38.3	44.6	36.6	13.3
	Innovative/control (unitless)	0.85	0.84	0.66	0.46
Sugar beet	Control system (t/ha)	99.3	105.1	93.1	73.4
	Innovative system (t/ha)	87.5	101.3	72.2	50.8
	Innovative/control (unitless)	0.88	0.96	0.78	0.69

**Fig. 5.** Lattice graph representing gross energy production for each cropping system type per year by platform in the Syppre experimental network. A point corresponds to a measurement on a plot. The lines connect the means. BEA: Béarn, BER: Berry, CHA: Champagne, LAU: Lauragais, PIC: Picardie, O: Innov: innovative system, X: control system, —: Evolution of the average of the innovative system, ---: Evolution of the average of the control system.

Lauragais was even worse than the control. Only the innovative system in Picardie showed an improvement compared to the control system. This shows that the overall reduction in TFI in innovative systems was partially compensated by the use of products with a higher concentration of active ingredients. Lastly, we observed a slight decrease over time of TFI (Fig. 7), an aspect that requires further monitoring for confirmation.

The results in terms of primary energy consumption and greenhouse gas emissions followed the same pattern with significantly improved performance for innovative systems.

These indicators are in fact strongly influenced by the amount of mineral nitrogen fertilizer used, which follows the same pattern (Fig. 8). The graphical analysis of GHG emissions shows that this overall result was valid for all platforms. This result is explained by the effect of (i) the introduction of legumes, and sometimes low-nitrogen consuming diversification crops such as sunflower or hemp in the crop sequences, and (ii) the optimization of nitrogen management strategies.

The energy efficiency of the innovative systems was not significantly different to that of the control systems. The high value of the *p*-value (Tab. 8) and the graphical analysis (Fig. 9)

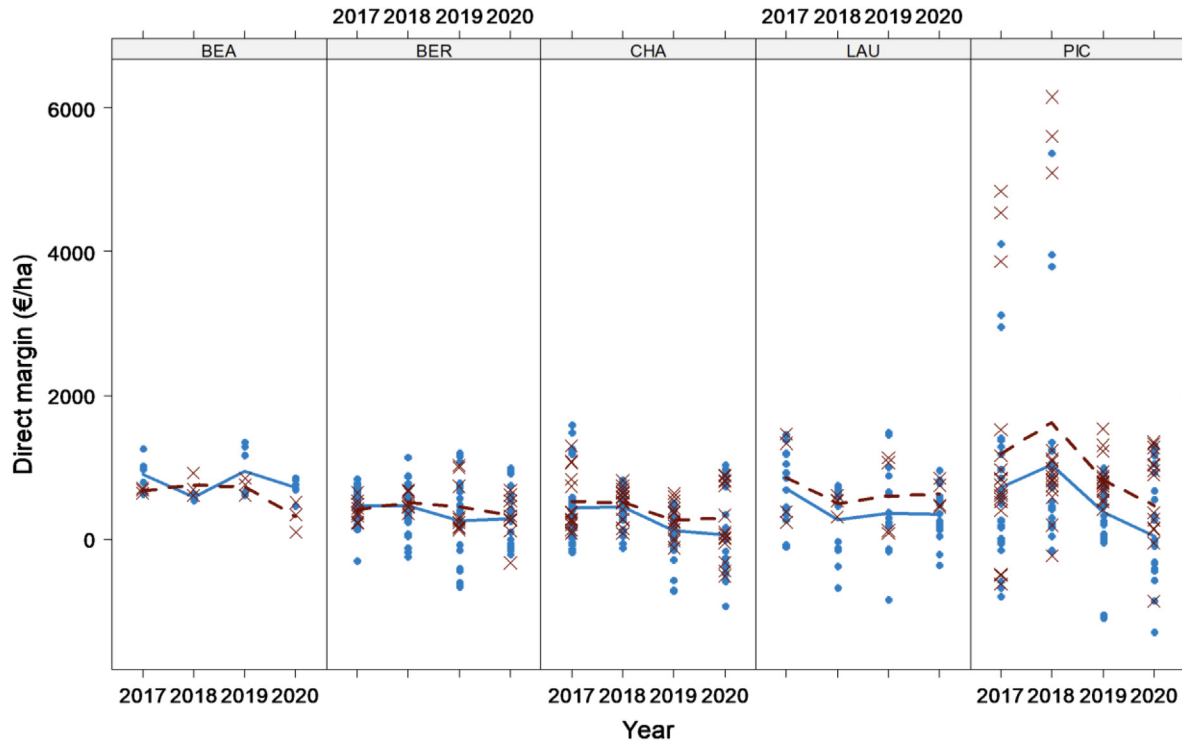


Fig. 6. Lattice graph representing direct margins with subsidies for each cropping system type per year by platform in the Syppre experimental network. A point corresponds to a measurement on a plot. The lines connect the mean values. BEA: Béarn; BER: Berry; CHA: Champagne; LAU: Lauragais; PIC: Picardie. O: Innov: innovative system, X: control system, — : Evolution of the average of the innovative system, ---: Evolution of the average of the control system.

suggest that the two types of systems were equivalent and that overall, the reduction in energy production was compensated by a reduction in energy consumption.

There was no significant difference between innovative and control systems for working time (Tab. 8). However, there were local specificities with a lower workload for the innovative system in Lauragais and a higher one in Béarn (Tab. 10).

Looking at all the performance criteria for each platform, we found that in the majority of the five platforms, the innovative systems performed better than the control systems in terms of environmental impact and input use but performed less well in terms of productivity and profitability (Tab. 10).

Are there antagonisms between indicators of performance?

On the whole, the considered indicators are moderately structured with an inertia of 33% on the first two factorial axes of the principal component analysis (Fig. 10). However, this allows us to highlight the indicators that are the most structuring. The analysis shows (i) a marked opposition between productivity indicators (negative values part of Dim 1, and frugality in inputs, especially primary energy use), and lower environmental impacts (positive values of Dim 1); (ii) correlations between economic performance and productivity indicators (the reduction in production costs does not seem to

compensate for the decline in production) and (iii) indicators that are independent of the others, such as energy efficiency, which do not appear in the first dimensions of the analysis (Fig. 10).

Dim 2 contrasts with the overall pattern observed for direct margins and the quantity of mineral nitrogen. This would reflect the lower profitability of the systems that integrate more legumes and other 'low-input' crops ('profitable' systems at the top, 'diversified with legumes and low-input crops' systems at the bottom). The correlation circle crossing dimensions one and four (not shown) allows the TFI indicator to appear with a positive value on Dimension 1, in contrast to the productivity indicators.

Do diversification and the cultivation of oilseed and protein crops result in improved cropping system performance?

The analysis of indicators at the block level was globally consistent with that conducted at the individual plot level. The first dimension highlighted a slope that placed the most "productive" systems towards the negative values of Dim 1 (Figs. 11 and 12) and the most "extensive" systems on the side of the positive values. In addition to this slope, two other trends were visible on Dim1: the higher proportion of oilseeds with positive coordinates and the distribution of platforms. The projection of performance indicators on the plane of the first two components of the PCA revealed that performance of

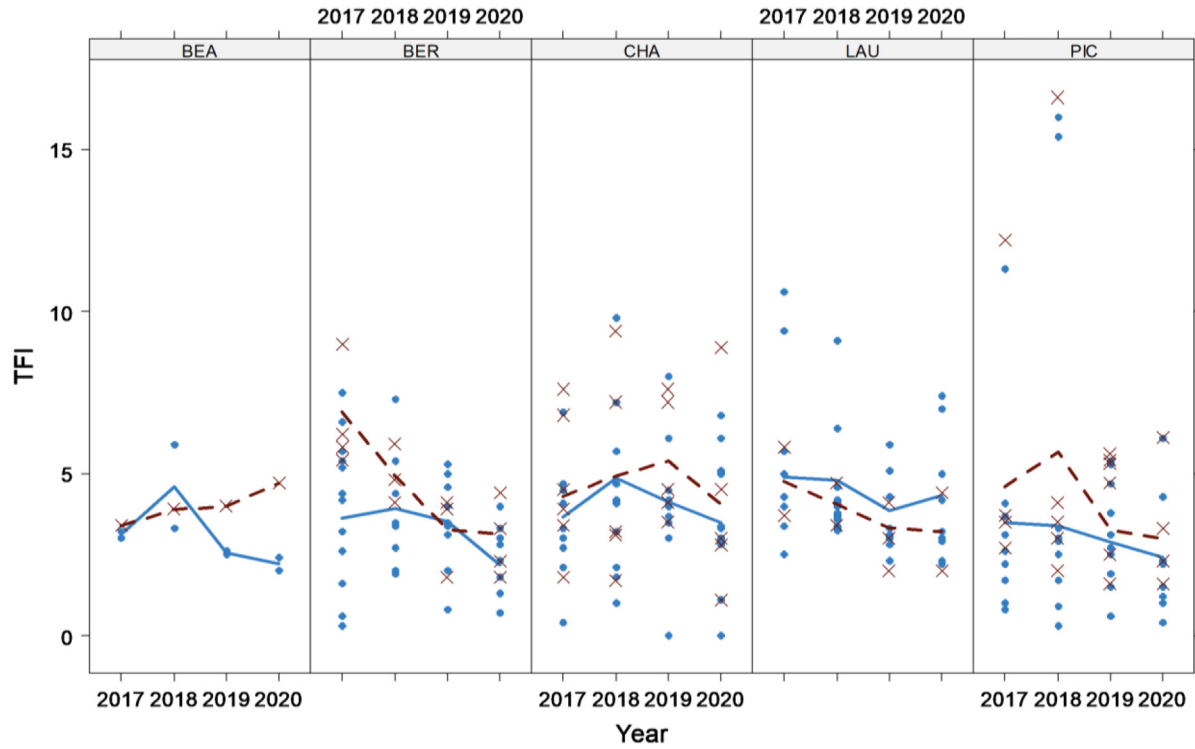


Fig. 7. Lattice graph representing treatment frequency Index for each cropping system type per year per platform in the Syppre experimental network. A point corresponds to a measurement on a plot. The lines connect the mean values. BEA: Béarn; BER: Berry; CHA: Champagne; LAU: Lauragais; PIC: Picardie. *O*: Innov: innovative system, *X*: control system, —: Evolution of the average of the innovative system, ---: Evolution of the average of the control system.

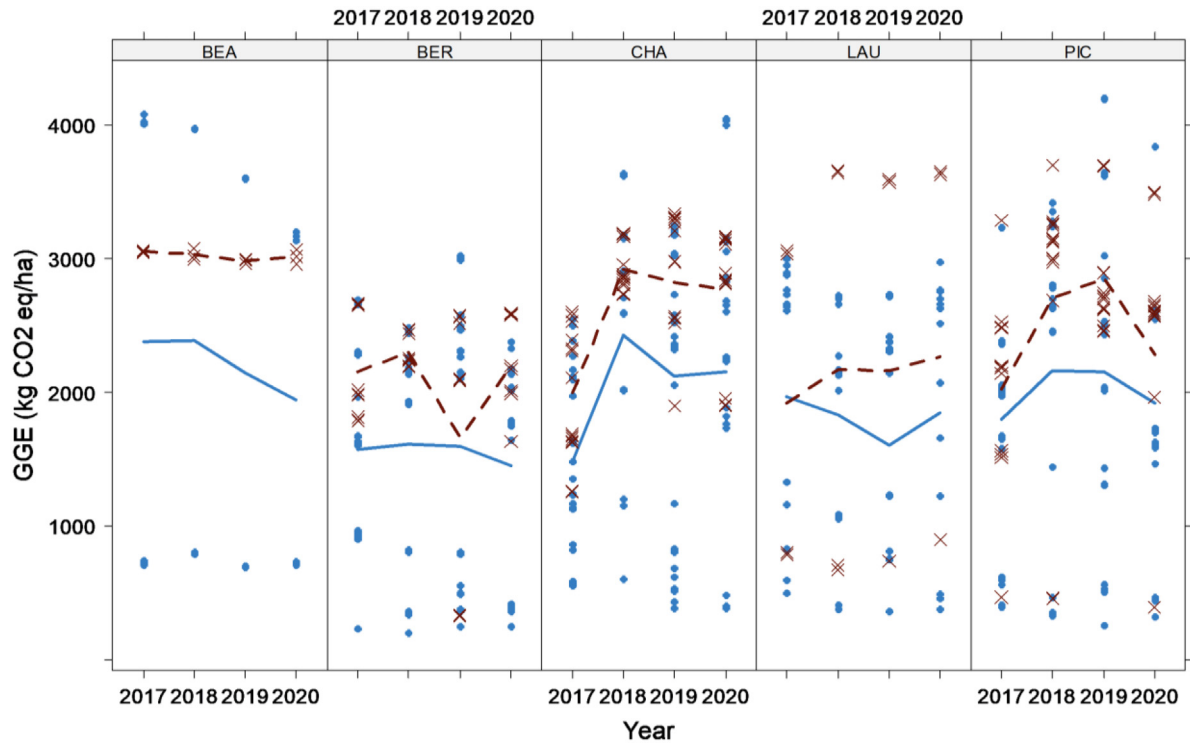


Fig. 8. Lattice graph representing total greenhouse gas emissions for each cropping system type per year per platform in the Syppre experimental network. A point corresponds to a measurement on a plot. The lines connect the mean values. BEA: Béarn; BER: Berry; CHA: Champagne; LAU: Lauragais; PIC: Picardie. *O*: Innov: innovative system, *X*: control system, —: Evolution of the average of the innovative system, ---: Evolution of the average of the control system.

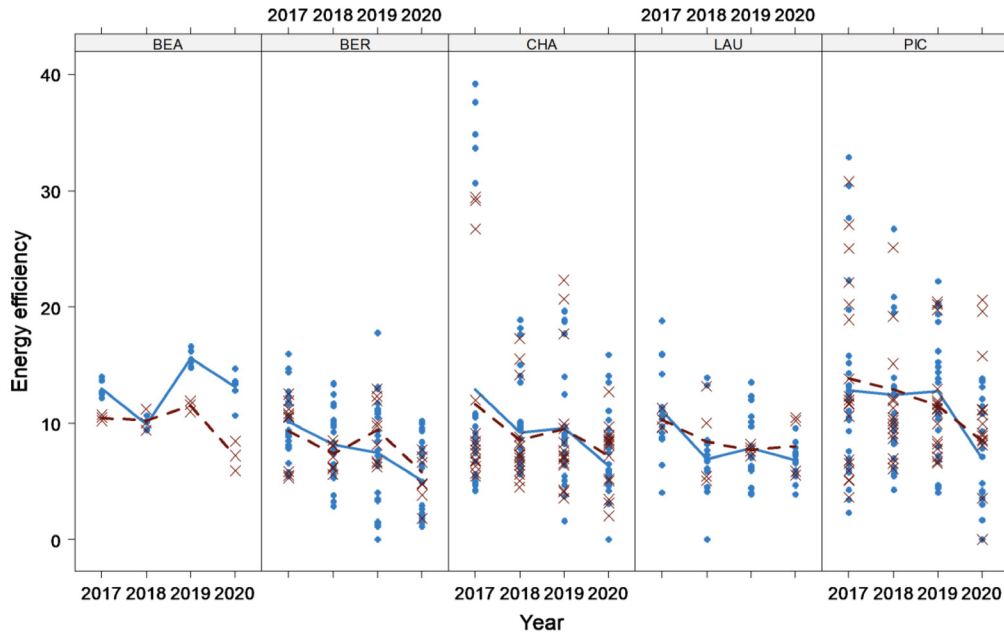


Fig. 9. Lattice graph representing energy efficiency for each system type per year per platform in the Syppre experimental network. A point corresponds to a measurement on a plot. The lines connect the means. BEA: Béarn; BER: Berry; CHA: Champagne; LAU: Lauragais; PIC: Picardie. *O*: Innov: innovative system, *X*: control system, —: Evolution of the average of the innovative system, ---: Evolution of the average of the control system.

Table 10. Qualitative assessment of the performance gap trend of the innovative cropping system compared to the control system for each of the five experimental platforms. The symbols mean that the results of the innovative system are inferior (<), inferior or equal (≤), equivalent (≈), superior or equal (≥) or superior (>) to the control system. The color code refers to the level of satisfaction, which refers to the absence of deterioration in economic performance, productivity and working time, and to the improvement in environmental performance and input use. Green means satisfied, red dissatisfied, yellow partially satisfied. BEA: Béarn; BER: Berry; CHA: Champagne; LAU: Lauragais; PIC: Picardie).

Indicator (y)	BEA	BER	CHA	LAU	PIC
Gross energy production	≈	<	<	≈	<
Energy efficiency	>	≈	≈	≈	≈
Direct margins with aids	≥	≈	≤	<	<
EBITDA per unit of human work time	≤	<	<	<	<
Total working time	>	≈	≈	<	≈
Treatment Frequency Index	<	<	<	>	<
Quantity of active pesticide ingredients	≈	≈	≈	>	<
Amount of mineral nitrogen fertilizer	<	<	<	<	<
Total primary energy used	<	<	<	≈	<
Total GHG emissions	<	<	<	<	<

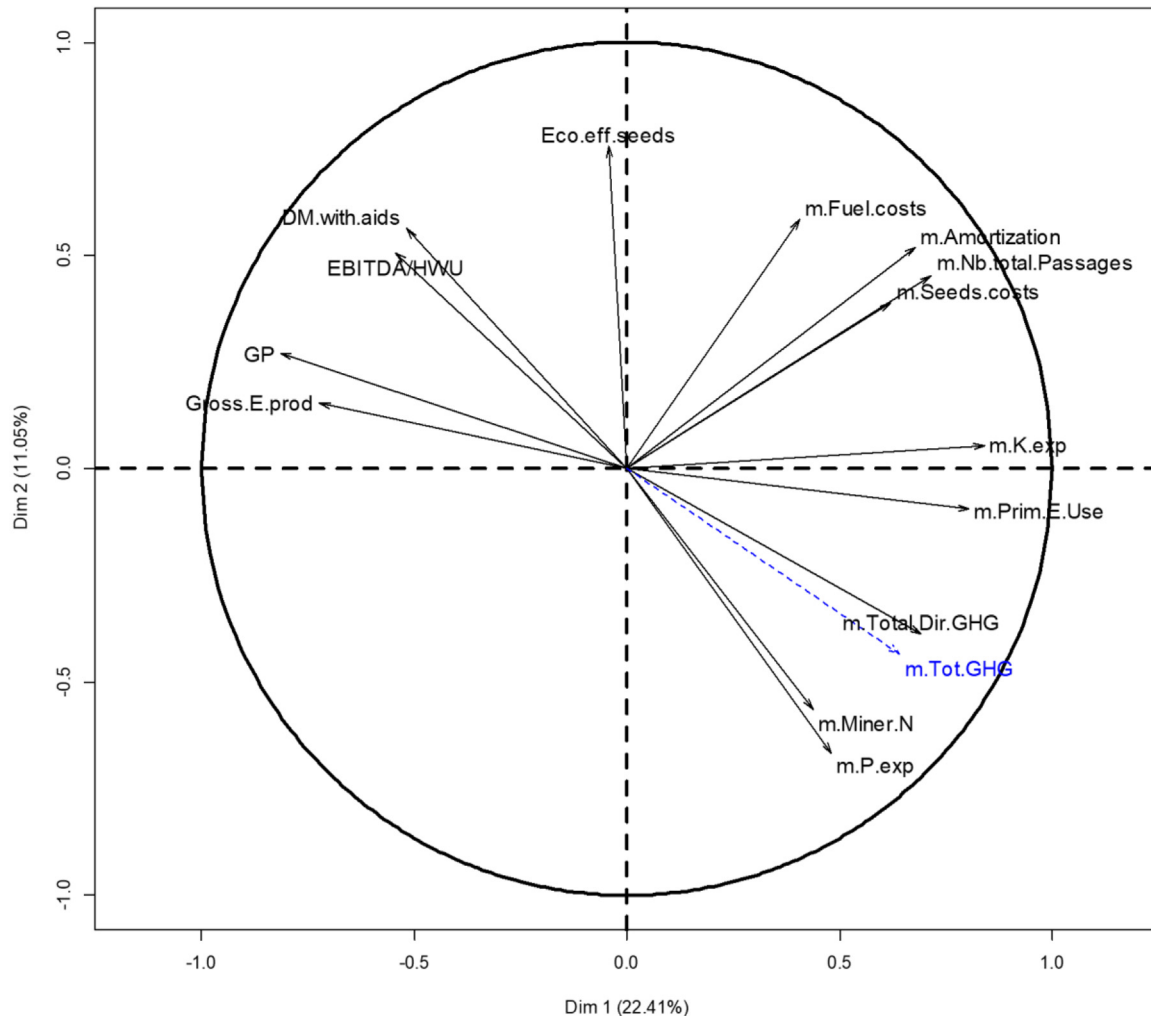


Fig. 10. Correlation circle crossing the first two components of the principal component analysis performed on the result performance indicators (53 in total) measured on plots. We represent here only the vectors whose $\cos^2$ (relative contribution) is greater than or equal to 50%. Indicators whose name begins with "m." are those whose value has been modified by adding a minus sign so that the direction of the vector indicates an improvement in the desired performance. Thus the arrow for "m.N_miner" indicates the direction of the decrease in nitrogen use (Tab. 6).

cropping systems was related more to production situations than the type of cropping system (innovative *versus* control; Fig. 12).

The analysis showed a slight difference between control and innovative systems. The cropping systems with the highest coordinates on Dim 2 were the innovative ones (Fig. 12). They were associated with higher proportions of legume crops, and therefore in principle with a lower amount of applied mineral nitrogen and lower greenhouse gas emissions, but also, more surprisingly, with a lower TFI (Fig. 11). This unexpected link could probably be explained by the fact that innovative systems, with more legume crops, also implement agroecological crop protection strategies. We also noted that there was no inversely correlated indicator on Dim 2, so there would not be any major opposition between the performance indicators of innovative cropping systems.

The clustering confirmed the major influence of the production situation on the performance profile and the minor influence of the type of cropping system (Fig. 13; see Appendix 2 for the descriptions of the obtained classes). With the exception of Béarn, both systems and all values of the same platform were always in the same cluster. The diversification descriptors did not have a marked impact on the performance profile. The platform in Béarn, a situation with high production potential, was notable in this ranking. An innovative system (Béarn I2) occupied a cluster alone (Fig. 13, cluster 1 in black). Its factorial position (negative coordinates on Dim 1 and positive coordinates on Dim 2) showed that this system would be one of those best reconciling multiple performance indicators. The description of this cluster, which was characterized by a higher proportion of protein crops, a lower proportion of oilseeds and a lower overall diversification,

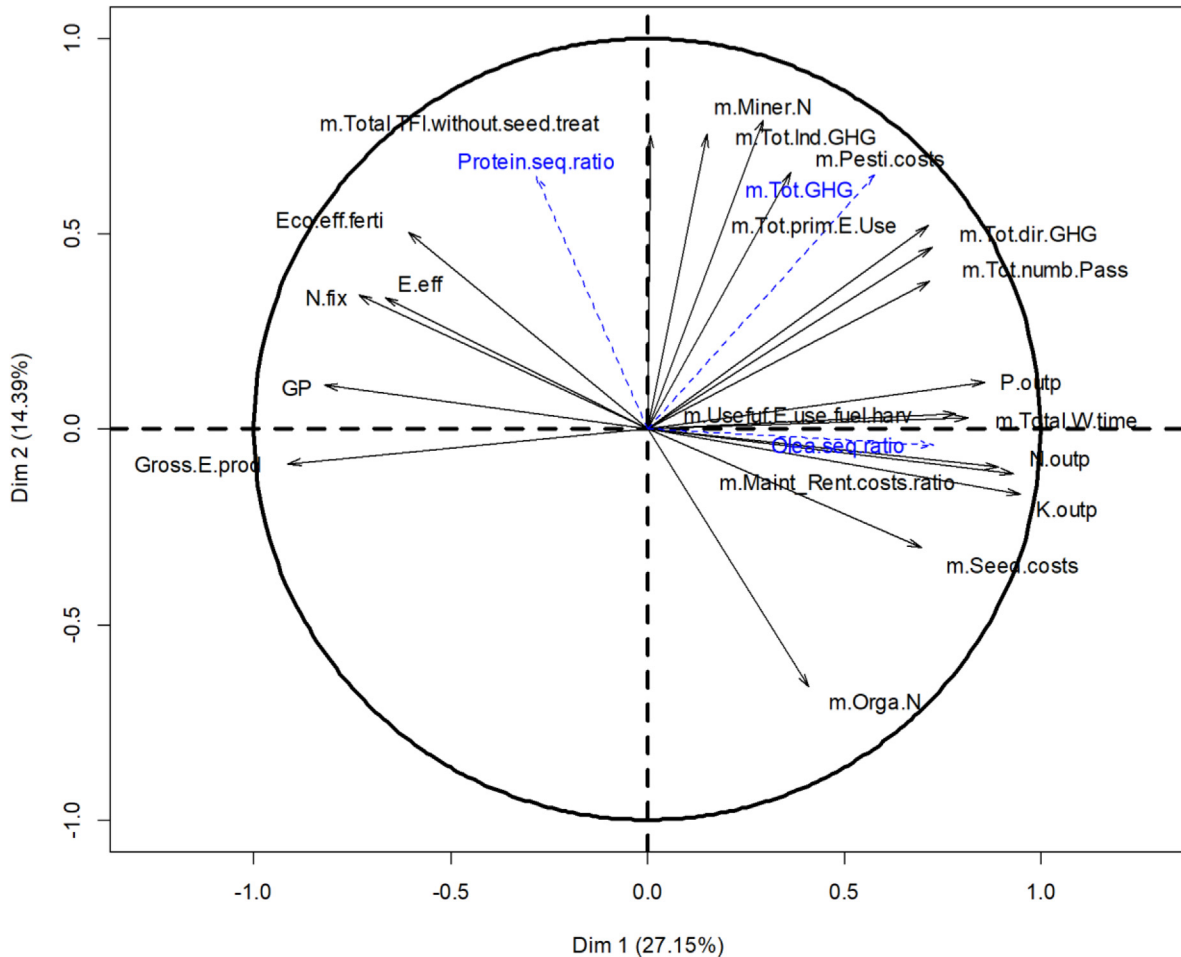


Fig. 11. Correlation circle crossing the first two components of the principal component analysis performed on the result indicators (53 in total) averaged by blocks. When indicators were redundant, only one was retained. We represent here only the vectors whose $\cos^2$ (relative contribution) is greater than or equal to 50%. Indicators whose name begins with "m." are those whose value has been modified by adding a minus sign so that the direction of the vector indicates an improvement in the desired performance. Thus, the arrow for "m.N_miner" indicates the direction of the decrease in nitrogen use. The blue vectors are illustrative and represent the indicators we have chosen to represent diversification as well as the oilseed and protein crop rates.

confirmed this assumption since it presented a high value of gross energy production, high energy efficiency, a high direct margin, and also low TFI and mineral nitrogen input.

Lessons learned and discussion

The analysis of data collected from the Syppre experimental network demonstrates the importance in system experiments of the interconnections between methodological considerations and the agronomic questions being addressed.

The planning of experiments is the key stage in answering agronomic questions

Beyond a non-exhaustive inventory of methodological issues, it was possible to identify their respective potential

impact on the experiments and the solutions to deal with them. Among the most noteworthy solutions, we showed that the planning of experiments is the key stage for avoiding certain pitfalls and ensuring that agronomic issues are properly addressed. Schillinger (2011) went as far even as to recommend “involve a statistician from the very first to ensure that the experimental design is valid and the most appropriate for the study”.

We have seen that under these conditions, experimental data can be processed with statistical tools such as mixed models and multifactorial analyses. As in other experimental domains, the main limiting factor is the size of the data set. It is therefore crucial that block-based experimental designs be used to enable spatial and temporal constraints to be managed, while contributing to the increase of the size of the collected data set.

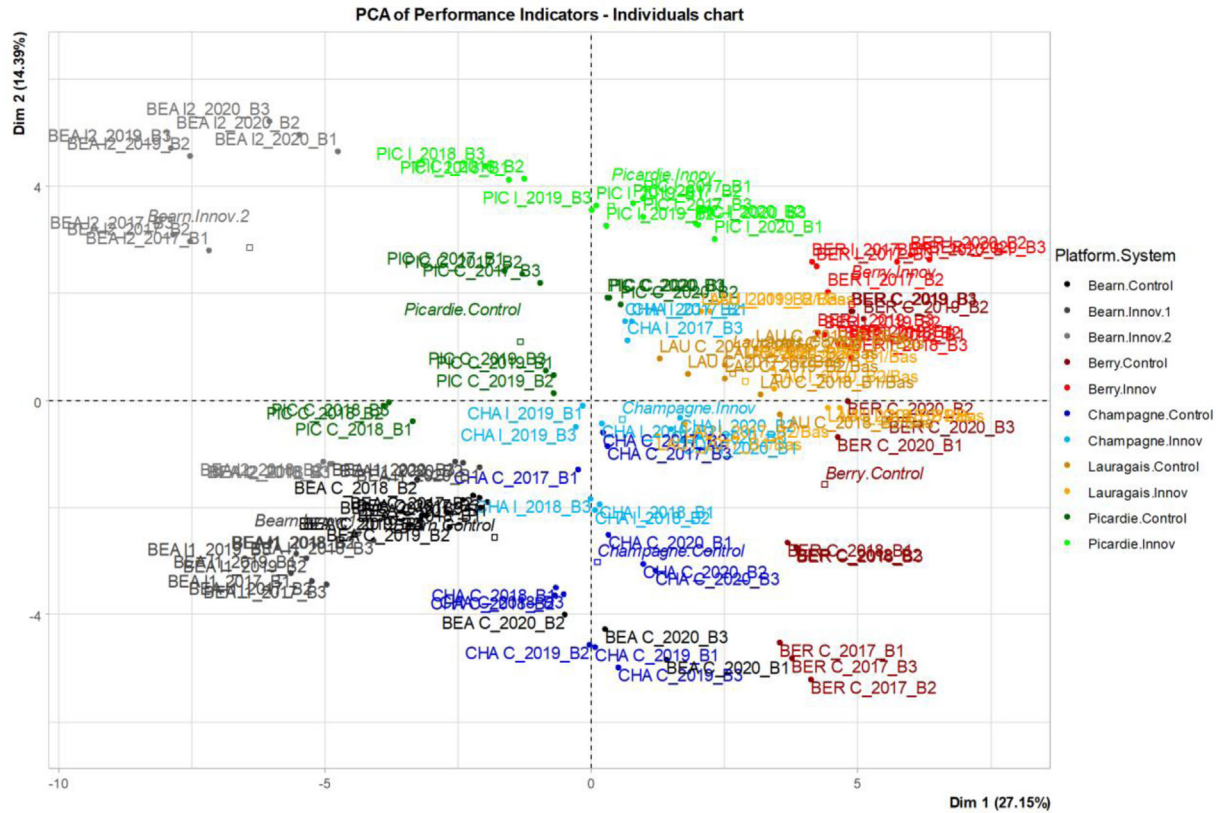


Fig. 12. Projection of individuals on the plan of the first two components of the principal component analysis of performance indicators by block. Each main color corresponds to a platform and the systems being tested there are identified by a variation of shade.

Experimenting in a network has many benefits

Our case study showed the benefits of working in the framework of an experimental network with coordinated platforms, with common objectives and standardized protocols. In such a situation, the heterogeneity brought by the platforms can be integrated in the planning phase so that analyses are not biased. Hence, while the size of the datasets from individual platforms was insufficient for drawing conclusions on general agronomic questions, not specific to platforms themselves, the entire dataset considered as a whole was indeed sufficient for statistical tools to be employed to address them.

We have also highlighted the value of having a control system in these experiments. Beyond the objective of assessing innovative cropping systems, the presence of a control system is essential for long-term studies because it enables cumulative system effects over time to be dissociated from pedoclimatic effects that apply to all of the cropping systems being tested. This appears especially important since the mixed models allowed us to conclude that the “year” effect often had a greater impact on the variation of indicators than the cropping system effect *per se*. The year effect might well be explained by meteorological and/or economic fluctuations, but information on these would have to be integrated into the dataset and analyzed for this to be confirmed.

Cropping systems and statistical units must be clearly defined

The statistical analyses could be completed only once two other related methodological questions had been solved: How to describe a cropping system in a cropping system experiment? What is the appropriate statistical unit to consider?

The first question had not been identified initially but was revealed during exchanges between experimenters and analysts in the course of the study. This highlights the importance of these regular exchanges all along the experimental project.

According to [Sebillotte \(1990\)](#), a cropping system is defined as a set of management operations implemented on identically cultivated plots (one or more fields; spatial dimension). Each cropping system is defined by the crop sequence, and the crop management system associated to each crop, including cultivar choice. The crop management system (“itinéraire technique” in French) has been defined as a logical and ordered combination of techniques which make it possible to control the environment and to derive a given production from it ([Sebillotte, 1974](#)). Cropping systems are thus paced by the cultivation of different crops (except for monocultures) and have effects on soil properties that cumulate over the years (temporal dimension). Due to lack of space, cropping systems

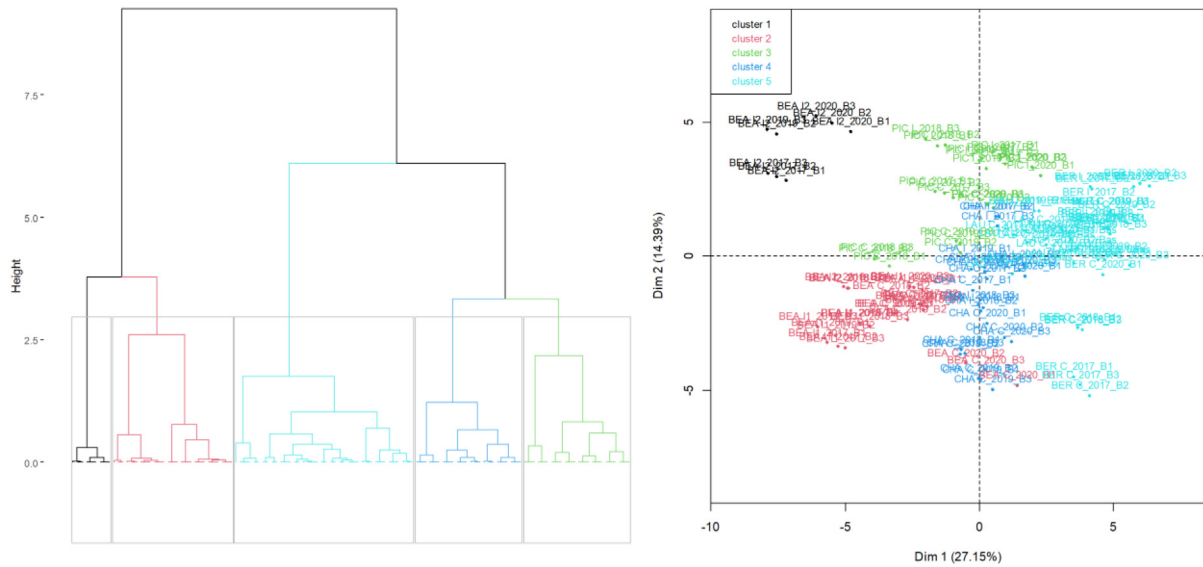


Fig. 13. Graphs resulting from the hierarchical ascending clustering performed on the PCA factorial coordinates using 53 result performance indicators. A partition in 5 classes was made. Description of the classes resulting from this clustering operation is provided in [Table S2](#).

are often tested solely in terms of their temporal dimension ([Lechenet *et al.*, 2017](#)), which requires experimenters to wait until the end of a rotation to conclude on the performance of the cropping systems being tested. For the Syppre experiments, the choice was made to consider both the temporal and spatial dimensions with the annual reproduction of all the crops in the rotation. In these circumstances, cropping systems can be considered in two ways: repeated on each plot with a different crop of the rotation each year; repeated on each block each year with all the crops of the rotation. In this case, the annual results of the crop sequence (spatially distributed crops of the rotation) are considered as an annual realization of the considered cropping system. In this approach, after short periods of time, the cumulative impacts of preceding crops and cropping practices are not well taken into account. However, this approach permits to provide useful information in an efficient way (*i.e.*, in a short time), at the cost of neglecting cumulative impacts of preceding crops, preceding cropping practices, and slow mechanisms. The longer the experiment, the better cumulative effects of cropping systems can reveal. Thus, there is a trade-off between time depth of experimental approaches and relevance of agronomic results which must be up to date for farmers.

From a methodological standpoint, the presence of all the crops of a rotation and the repetition of the crop rotation each year means that the rotation systems under test can be evaluated on both their temporal and spatial dimensions. Depending on the agronomic question being considered, it will be possible to use the "plot" (rotation over time) or the "block" (rotation within the plots) as the statistical unit. Thus, even if cropping systems have not yet completed their first rotation in time, it is still possible to draw annual conclusions and to study the first temporal evolutions and so the transition period during which the cumulative effects of new cropping system appear gradually. However, this early analysis does not allow to draw conclusions about long term processes such as weed infestation, or soil carbon sequestration. Definitive conclusions on performance of the tested cropping systems therefore require a longer

time, at least one rotation length. It was possible to show that using data per plot was more relevant for comparing cropping systems and for studying discrepancies between indicators. Also, the study of links between diversification and multiple performance indicators can be performed only by using blocks as statistical units. In addition, when the study of multidimensional performance involves diversification, this can only be conducted at the level of rotations. More generally, we believe that the analysis of data from a system experiment requires several data tables, the number depending on the statistical unit being considered. In order to save time and avoid mistakes, we recommend anticipating this multi-level aspect when constructing databases. Likewise, the storage tools must permit the simplified and personalized addition of new variables, such as diversity indices, that are essential for answering certain agronomic questions. To assess the implications of diversification and the introduction of oilseeds and protein crops, we had to conduct tedious data manipulation between two tables. To avoid this type of inconvenience, we recommend ensuring at the earliest stage that all the variables relevant for future analyses have been incorporated into the dataset's structure.

Many methodological issues remain to be resolved

We have chosen to present answers to methodological issues of primary importance. However, there are still many questions that merit being addressed. In particular, how to evaluate the "preceding crop effect?" How to deal with data for which measurements have not been carried out every year? Or how to deal with crop data when there are two harvests per year (a frequent situation when cover crops are cultivated for energy purposes).

Finally, it should be borne in mind that the methodological issues identified with the Syppre experimental network are common to many other system experiments. Also, there are other methodological issues that have not been addressed here due to the diversity of objectives, themes, domains or

experimental designs. For example, we did not include any questions about the analysis of biodiversity measurements, the evaluation of decision rules, or the analysis of cropping systems when not all crops of the rotation are present each year, or when there are no replicates.

Agronomic results

Reconciling multiple indicators of performance

There is a general consensus in the literature that agroecology is the way for reconciling multiple indicators of performance in production systems (Tilman *et al.*, 2002; Robertson *et al.*, 2005; Malézieux, 2012; Duru *et al.*, 2015). However, published results on concrete cases remain scarce. Our results show that developing an innovative system that reconciles multiple indicators of performance is not straightforward.

A first explanation is the antagonism between indicators of performance that was highlighted in our study. Even though we found a pattern of positive correlations between productivity and profitability criteria, a strong antagonism between productivity and primary energy consumption indicators was observed. This antagonism between productivity/profitability and input use/environmental impacts has been highlighted already in other studies (Rosa-Schleich *et al.*, 2019), notably in the French context (Bonnet *et al.*, 2021). In their study, Lechenet *et al.* (2014) did not detect any correlation between the intensity of pesticide use and either productivity or profitability. However, they did show that the productivity of integrated systems was lower than that of conventional systems, confirming our results which indicate the difficulty of maintaining productivity at the same level in agroecological systems after four years of transition. Regarding pesticide use, our study revealed unexpected results. While TFI were reduced in four of the five innovative systems, it was higher in one of them. This case of Lauragais was explained by a control system based on a rotation of two crops including sunflower, which is one of the arable crops with the lowest TFI in France (SSP – Agreste, 2019). The lengthening of the rotation to eight years, with five new crops having a higher TFI was not compensated by a reduction in the overall pest pressure and led to an increase in the TFI at rotation level. This result underlines the importance of considering local production conditions when explaining performance and the capacity of systems to reconcile multiple indicators of performance, as already highlighted by Beillouin *et al.* (2019) and Duru *et al.* (2015). The fact that the performance indicator clustering distinguished more the effects of production situations than the types of cropping systems strengthened this conclusion and should cause caution when generalizing the performance results of agroecological systems. This finding also suggests that deeper changes in local agri-food systems might help remove obstacles to more profound diversifications in cropping systems.

Another explanation for the difficulty of reconciling all the objectives is the lack of hindsight in the experiment, as the analysis only covers the first four years of experimentation. Indeed, innovative systems require time for technical learning (Colnenne-David *et al.*, 2017) and the effects of strategies

implemented to suppress weeds, animal pests and diseases are only gradual (Deguine *et al.*, 2017; Liu *et al.*, 2022). The case study revealed the difficulties in mastering innovative techniques, such as in Picardie with the planting of potatoes and sowing of sugar beet without plowing, or more generally in ensuring the success of lesser-known diversification crops. Rosa-Schleich *et al.* (2019) showed that the trade-off between ecological and economic indicators of performance was frequent in the short term and reported many examples of improved productivity and profitability over the longer term. However, with the lack of interaction between the cropping system and the year effects in our results, we could not show any progressive improvement effect of the innovative systems, as compared to the well-mastered control systems. If this effect had indeed existed, it would have had to be larger, in view of the weather and biotic stress hazards from one year to the next, if it was to be revealed within this time frame.

Future data will be useful to verify this hypothesis of progressive improvement and possibly to specify the period of time that has to lapse before a new balance is reached. Furthermore, economic performance should be analyzed to take into account: (i) their variability; (ii) a wider range of economic drivers, not only input prices, agricultural product prices, labor costs, but also the costs of all negative externalities (Bourguet and Guillemaud, 2016). In any case, this additional example of ecological-economic trade-off highlights the importance of considering the use of financial instruments to recognize environmental performance and/or to support the implementation of cropping system diversification, such as proposed by Rosa-Schleich *et al.* (2019).

Finally, we found that our *ex-post* results differed significantly from the results obtained through the *ex-ante* evaluation (Viguié *et al.*, 2021). In particular, the *ex-ante* evaluation showed an economic benefit for all five innovative systems, which was not observed in this study. This difference should be explained, as previously detailed, by the time needed to attain mastery of innovative cropping systems and benefits from their cumulative effects, as well as by a tendency to overestimate the performance of diversification crops in the *ex-ante* hypotheses, as observed by Colnenne-David *et al.* (2017).

Impact of diversification strategies on performance

The innovative cropping systems being tested were very diversified in comparison to those in other studies (*e.g.*, Bonnet *et al.*, 2021) and relative to the control systems. While diversification is often considered as a major way to reconcile multiple indicators of performance (Lin, 2011; Ratnadass *et al.*, 2012; Gaba *et al.*, 2014), our results were not as definitive.

Principal component analysis did not link the level of overall diversification to a particular performance profile, thereby indicating that diversification was neither favorable nor unfavorable to multiple indicators of performance. However, this absence of a link was partly explained by the major impact of the production situation in performance profiles. In spite of this, the cluster that contained only an innovative system reconciling multiple indicators of performance was characterized by a lower-than-average level of diversification. Furthermore, the detailed platform-by-platform results, which allowed us to overcome the production context effect, showed that, across all the

performance criteria, the more diversified innovative cropping systems showed no improvement compared to the less diversified control cropping systems, especially in terms of productivity and profitability. The example of Berry illustrates this point very well: the profitability of the innovative system was better than that of the control in 2017, a year without major weather hazards, but worse in the three following years when severe droughts were observed in the spring and/or summer (Fig. 5), contrary to the common impression (Lin, 2011). In this case, the diversification of the system with spring crops to improve weed control rendered the system less robust in the face of climatic events. These results showed that it was not by diversifying to the maximum that performance was enhanced, but rather by diversifying with complementary crops, and by finding the correct balance with the introduction of diversification crops best suited to the production situation. This question of balance in the diversification of cropping systems has been the subject of only a small amount of research. Zampieri *et al.* (2020) found that at country-level scale (France), production resilience increased with crop diversity but levelled-off at six crops. The question was more intensively addressed for cover crops. Numerous studies have failed to show any benefit of species diversification in cover crops to the average performance of a range of criteria, or even on the stability of that performance (Florence *et al.*, 2019; Florence and McGuire, 2020; Smith *et al.*, 2020). These studies showed that the most diverse mixtures (more than 5 to 10 species) were never the best performing. Smith *et al.* (2020) explained these results by the fact that the more diverse the mixture, the more the share of the best performing species is reduced, at the expense of the overall performance of the mixture. In our study, this dilution effect could outweigh, at least in the short term, the positive "system effect" of crop diversification (*e.g.*, related to improved pest control, or soil fertility) and thus explained the mixed performance of highly diversified systems as in Berry and Lauragais. The question of how the level of crop diversification might be adapted according to the production situation merits additional research effort.

Regarding the impact of legume crops, our results showed that increasing their proportion in the rotation was correlated with a lower consumption of mineral nitrogen and, consequently, a lower consumption of primary energy and lower greenhouse gas emissions at the scale of the rotation. This confirms the results from numerous studies on the benefits related to cropping systems that result from the introduction of legumes (Drinkwater *et al.*, 1998; Jensen *et al.*, 2012; Reckling *et al.*, 2016; Liu *et al.*, 2022). More surprisingly, our results showed a correlation between the proportion of legumes in the crop rotation and low TFI. However, we suggest that this was not a causal relationship but: (i) a confusion of effects knowing that the innovative systems had both an objective of reducing nitrogen inputs and associated impacts (all integrated more legumes) and also TFI (all integrated strategies to reduce the use of pesticides); (ii) the weight of the Béarn's innovative system with a very high proportion of legumes (soybean, one of the field crops with the lowest TFI in France, Agreste, 2019), which had a strong influence on the overall results. For the other performance indicators, we did not find any trade-off with the proportion of legumes, thus confirming the essential role of legumes in cropping systems' performance.

The impact of oilseed crop proportion could not be interpreted in our study due to the strong effect of the production situation.

Our results highlight the importance of continuing to generate knowledge on agroecological strategies, taking into account local specificities, and avoiding making generalizations about unspecific trends. In particular, it is important to better understand how the level of diversification can best be balanced and to verify whether short-term performance trade-offs fade over the longer term. In addition, many authors argue that the production and diffusion of general scientific knowledge is essential but by itself not enough to support agroecological transitions (Duru *et al.*, 2015). There is a need to develop support approaches based on scientific knowledge to support farmers in the re-design of tailor-made systems adapted to their own situations (Le Gal *et al.*, 2011; Duru *et al.*, 2015). Developing co-innovation approaches and hybridizing on-station experiments and on-farm design support are other objectives of the Syppre project (Cadoux *et al.*, 2019) whose results will contribute to supporting agroecological transition.

Acknowledgements

The authors sincerely thank all the engineers, technicians and interns involved in the Syppre experimental network. They thank Jean Villerd (INRAE) for his role in the supervision of the PhD of the first author. They are also grateful to Virginie Brun (APCA), Marie Coquet Rougier (ACTA), Florent Duyme (Arvalis-Institut du Végétal), Emeric Emonet (ACTA), and Marie-Sophie Petit (Agricultural Chamber of Burgundy) for their useful advice, as members of the PhD steering committee of the first author. The authors thank the anonymous referees and the editor for their helpful comments that improved the quality of the manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

Funding

Syppre is part of the European project 'Diversification through Rotation, Intercropping, Multiple Cropping, Promoted with Actors and value-Chains towards Sustainability' (Diver-IMPACTS), funded by the European Commission under Grant Agreement number 727482. This work was also partially funded by the French Ministry of Agriculture and Food through the special allocation account for agricultural and rural development CASDAR. The PhD thesis of the first author was funded by Arvalis-Institut du Végétal and the French National Research Agency.

Supplementary Material

Table S1. List of indicators used in the analysis of the Syppre experimental network data as well as the course names displayed in the multi-factor graphs. 1: indicators used for mixed models. 2: indicators used as active variables in PCA.

Table S2. Description of the classes resulting from clustering.

The Supplementary Material is available at <http://www.ocl-journal.org/10.1051/ocl/2023022/olm>.

References

- Agreste. 2019. Pratiques culturales en grandes cultures 2017. 30p.
- Aubertot JN, Robin MH. 2013. Injury profile SIMulator, a qualitative aggregative modelling framework to predict crop injury profile as a function of cropping practices, and the abiotic and biotic environment. I. conceptual bases. *PLoS One* 8.9: e73202. <https://doi.org/10.1371/journal.pone.0073202>.
- Baayen RH, Davidson DJ, Bates DM. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang* 59: 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>.
- Beillouin D, Ben-Ari T, Makowski D. 2019. A dataset of meta-analyses on crop diversification at the global scale. *Data in brief* 24: 103898. <https://doi.org/10.1016/j.dib.2019.103898>.
- Bianconi A, Dalgaard T, Manly BFI *et al.*, 2013. Methodological difficulties of conducting agroecological studies from a statistical perspective. *Agroecol Sustain Food Syst* 37: 485–506. <https://doi.org/10.1080/10440046.2012.712941>.
- Boiffin J, Malezieux E, Picard D. Cropping systems for the future. In Nösberger J, Geiger HH, Struik PC, eds. Crop science: Progress and prospects. Papers presented at the Third International Crop Science Congress, Hamburg, Germany 17–22 August 2000. CABI Publishing, 2001, pp. 261–279.
- Bonnet C, Gaudio N, Alletto L, *et al.*, 2021. Design and multicriteria assessment of low-input cropping systems based on plant diversification in southwestern France. *Agron Sustain Dev* 41. <https://doi.org/10.1007/s13593-021-00719-7>.
- Bourguet D, Guillemaud T. 2016. The hidden and external costs of pesticide use. In: Lichtfouse, E. (eds). Sustainable Agriculture Reviews, vol 19. Springer, Cham. https://doi.org/10.1007/978-3-319-26777-7_2
- Breman H, de Wit CT. 1983. Rangeland productivity and exploitation in the Sahel. *Science* 221.4618: 1341–1347. <https://doi.org/10.1126/science.221.4618.1341>.
- Brossier J, Contini C, Zorini LO, Cristóvão A. The origins of the European IFSA : the first meetings and the agenda renewal. In I. Darnhofer I., D. Gibbon D, eds. Farming Systems Research into the 21st Century: The New Dynamic. Springer Science & Business Media 2012, p 33–48.
- Brühl CA, Zaller JG. 2019. Biodiversity Decline as a Consequence of an Inappropriate Environmental Risk Assessment of Pesticides. *Front Environ Sci* 7: 177. <https://doi.org/10.3389/fenvs.2019.00177>.
- Cadoux S, Sauzet G, Pruvot T, Flénet F. Diversification of cropping systems in a context of shallow soils : implementation and performances of both on-station and on-farm systems. In Messéan A, Drexler D, Heim I, Paresys L, Stilmant D, Willer H, eds. European Conference on Crop Diversification. INRA and ÖMKI Budapest, Hungary, 2019, pp. 39–40.
- Cady FB. 1991. Experimental design and data management of rotation experiments. *Agron J* 83: 50–56. <https://doi.org/10.2134/agronj1991.00021962008300010014x>.
- Colnenne-David C, Grandeau G, Jeuffroy MH, Dore T. 2017. Ambitious environmental and economic goals for the future of agriculture are unequally achieved by innovative cropping systems. *Field Crops Research* 210: 114–128. <https://doi.org/10.1016/j.fcr.2017.05.009>.
- Cros MJ, Aubertot JN, Gaba S, Reboud X, Sabbadin R, Peyrard N. 2021. Improving pest monitoring networks using a simulation-based approach to contribute to pesticide reduction. *Theor Popul Biol* 141: 24–33. <https://doi.org/10.1016/J.TPB.2021.06.002>.
- Dagnelie P. 2012. Principes d'expérimentation. Planification des expériences et analyse de leurs résultats. *Gembloux, Belgique: Les Presses agronomiques de Gembloux, A.S.B.L.*
- De'ath G. 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* 83: 1105–1117.
- Debaeke P, Munier-Jolain N, Bertrand M, Guichard L, Nolot JM, Saulas P. 2009. Agronomy for Sustainable Development 29 (1) 73–86. 2009. Iterative design and evaluation of rule-based cropping systems: methodology and case studies. A review. <https://doi.org/10.1051/agro:2008050>
- De Cordoue A-L., Toqué C, Cadoux S, *et al.* Syppre: synergy between experimental platforms, groups of farmers and observatory of current farming systems to support transition towards new cropping systems. In: *Book of abstracts of the XV European Society for Agronomy Congress: "Innovative cropping and farming systems for high quality food production systems"*, Genève: Agroscope, 2018.
- De Cordoue A, Cadoux S, Toqué C, *et al.* Co-design OF agro-ecological cropping systems combining global and local issues. In: *Book of abstracts of the XIV European Society for Agronomy Congress: "Growing landscapes –Cultivating innovative agricultural systems"*, Edinburgh, Scotland, 2016.
- Deguine JP, Gloanec C, Laurent P, Ratnadass A, Aubertot, JN. Agroecological crop protection, *Springer*, 2017.
- Doré T, Bail M Le, Martin P, Ney B, Roger-Estrade J. 2006. L'agronomie aujourd'hui, *Ed. Quae*.
- Drinkwater LE, Wagoner P, Sarrantonio M. 1998. Legume-based cropping systems have reduced carbon and nitrogen losses. *Nature* 396.6708: 262–265.
- Dubois S, Tailleur A, Cadoux S, *et al.* Evaluation multicritère a priori pour aider à la conception de systèmes Syppre: Répondre aux défis de l'agriculture Concilier agronomie et écologie pour produire. In Colbach N, Angevin F, Bockstaller C, *et al.*, eds. Projet COSAC Colloque final. hal-02791468, Paris, France 2019, pp. 36–38.
- Duchene O, Vian JF, Celette F. 2017. Intercropping with legume for agroecological cropping systems: Complementarity and facilitation processes and the importance of soil microorganisms. A review. *Agric Ecosyst Environ* 240: 148–161. <https://doi.org/10.1016/J.AGEE.2017.02.019>.
- Duru M, Therond O, Martin G, *et al.*, 2015. How to implement biodiversity-based agriculture to enhance ecosystem services: a review. *Agron Sustain Dev* 35: 1259–1281. <https://doi.org/10.1007/s13593-015-0306-1>.
- Elhakeem A, van der Werf F W, Bastiaans L. 2021. Radiation interception and radiation use efficiency in mixtures of winter cover crops. *Field Crops Research* 264: 108034. <https://doi.org/10.1016/j.fcr.2020.108034>.
- FAO. 2019. Sustainable food and agriculture: an integrated approach, *Academic Press*.
- Federer WT. 1999. Statistical design and analysis for intercropping experiments volume II: three or more crops, New York: Springer.
- Florence AM, Higley LG, Drijber RA, Francis CA, Lindquist JL. 2019. Cover crop mixture diversity, biomass productivity, weed suppression, and stability. *PLoS One* 1–18. <https://doi.org/10.1371/journal.pone.0206195>.
- Florence AM, McGuire AM. 2020. Do diverse cover crop mixtures perform better than monocultures? A systematic review. *Agron J* 112: 3513–3534. <https://doi.org/10.1002/agj2.20340>.

- Foley JA, Ramankutty N, Brauman KA, *et al.*, 2011. Solutions for a cultivated planet. *Nature* 478: 337–342. <https://doi.org/10.1038/nature10452>.
- Gaba S, Lescourret F, Boudsocq S, *et al.*, 2014. Multiple cropping systems as drivers for providing multiple ecosystem services: from concepts to design. *Agron Sustain Dev* 35.2: 607–623. <https://doi.org/10.1007/s13593-014-0272-z>.
- Gac A, Deltour L, Cariolle M, *et al.*, 2010. GES'TIM. Guide méthodologique pour l'estimation des impacts des activités agricoles sur l'effet de serre. *Version 1. 2*.
- Gac A, Cariolle M, Deltour L, *et al.*, 2011. GES'TIM – des apports pour l'évaluation environnementale des activités agricoles. *Innovations Agronomiques, INRAE* 17: 83–94.
- Godfray HCJ, Beddington JR, Crute IR, *et al.*, 2010. Food security: the challenge of feeding 9 billion people. *Science* 327.5967: 812–819. <https://doi.org/10.1126/science.1185383>.
- Gomez KA, Gomez AA. 1984. Statistical procedure for agricultural research, 2nd ed. New York: John Wiley & Sons.
- Hallmann CA, Sorg M, Jongejans E, *et al.*, 2017. **More than 75 percent decline over 27 years in total flying insect biomass in protected areas.** *PLoS One* 12.10: e0185809. <https://doi.org/10.1371/journal.pone.0185809>.
- IPCC. 2006. Guidelines for national greenhouse gas inventories. *Volume 4. Agriculture, Forestry and Other Land Use*.
- IPCC. 2019. Climate Change and Land. Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. *Summary for Policymakers. Geneve*.
- Jensen ES, Peoples MB, Boddey RM, *et al.*, 2012. Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. *A review. Agron Sustain Dev* 32.2: 329–364. <https://doi.org/10.1007/s13593-011-0056-7>.
- Joly D, Brossard T, Cardot H, Cavailles J, Hilal M, Wavresky P. 2010. Les types de climats en France, une construction spatiale. *CyberGeo : Eur J Geogr* 501: 1–25. <http://journals.openedition.org/cybergeogeo/23155> <https://doi.org/10.4000/cybergeogeo.2315>.
- Keichinger O, Viguier L, Corre-Hellou G, Messean A, Angevin F, Bockstaller C. 2021. Un indicateur évaluant la diversité globale des rotations: de la diversité des cultures aux services écosystémiques. *AE&S* 11-1: 1–14.
- Kudsk P, Jørgensen LN, Ørum JE. 2018. Pesticide load—a new Danish pesticide risk indicator with multiple applications. *Land Use Policy* 70: 384–393. <https://doi.org/10.1016/j.landusepol.2017.11.010>
- Lamanda N, Roux S, Delmotte S, *et al.*, 2012. A protocol for the conceptualization of an agro-ecosystem to guide data acquisition and analysis and expert knowledge integration. *Eur J Agron* 38: 104–116. <https://doi.org/10.1016/j.eja.2011.07.004>.
- Laurent P, Aubertot JN, Doré T, *et al.*, 2017. Agroecological transition keys. In Deguine JP, Gloanec C, Laurent P, Ratnadass A, Aubertot, JN, eds. *Agroecological Crop Protection*. Springer, Dordrecht 2017, pp. 163–246.
- Le Gal PY, Dugué P, Faure G, Novak S. 2011. How does research address the design of innovative agricultural production systems at the farm level? A review. *Agric Syst* 104: 714–728. <https://doi.org/10.1016/j.agsy.2011.07.007>.
- Lechenet M, Bretagnolle V, Bockstaller C, *et al.*, 2014. Reconciling pesticide reduction with economic and environmental sustainability in arable farming. *PLoS One* 9.6: e97922. <https://doi.org/10.1371/journal.pone.0097922>.
- Lechenet M, Deytieux V, Antichi D, *et al.*, 2017. Diversity of methodologies to experiment Integrated Pest Management in arable cropping systems: analysis and reflections based on a European network. *Eur J Agron* 83: 86–99.
- Lescourret F, Dutoit T, Rey F, Côte F, Hamelin M, Lichtfouse E. 2015. Agroecological engineering. *Agron Sustain Dev* 35: 1191–1198. <https://doi.org/10.1007/s13593-015-0335-9>.
- Lin BB. 2011. Resilience in agriculture through crop diversification: adaptive management for environmental change. *Bioscience* 61: 183–193. <https://doi.org/10.1525/bio.2011.61.3.4>.
- Liu C, Plaza-Bonilla D, Coulter JA, *et al.*, 2022. Diversifying crop rotations enhances agroecosystem services and resilience. *Adv Agron* 1–37. <https://doi.org/10.1016/bs.agron.2022.02.007>.
- Malézieux E. 2012. Designing cropping systems from nature. *Agron Sustain Dev* 32: 15–29. <https://doi.org/10.1007/s13593-011-0027-z>.
- Martin G, Martin-Clouaire R, Duru M. 2013. Farming system design to feed the changing world. A review. *Agron Sustain Dev* 33: 131–149. <https://doi.org/10.1007/s13593-011-0075-4>.
- Meynard JM, Dedieu B, Bos AP. Re-design and co-design of farming systems. An overview of methods and practices. In Darnhofer I, Gibbon D, Dedieu B, eds. *Farming Systems Research into the 21st Century: The New Dynamic*. Springer Science & Business Media, 2012, pp. 405–429.
- Plaza-Bonilla D, Nolot JM, Raffaillac D, Justes E. 2017. Innovative cropping systems to reduce N inputs and maintain wheat yields by inserting grain legumes and cover crops in southwestern France. *Eur J Agron* 82: 331–341. <https://doi.org/10.1016/j.eja.2016.05.010>.
- R Core Team. 2020. *R: A language and environment for statistical computing*.
- Ratnadass A, Fernandes P, Avelino J, Habib R. 2012. Plant species diversity for sustainable management of crop pests and diseases in agroecosystems: A review. *Agron Sustain Dev* 32 : 273–303. <https://doi.org/10.1007/s13593-011-0022-4>.
- Reckling M, Bergkvist G, Watson CA, *et al.*, 2016. Trade-offs between economic and environmental impacts of introducing legumes into cropping systems. *Front Plant Sci* 7: 669. <https://doi.org/10.3389/fpls.2016.00669>.
- Robertson GP, Swinton SM. 2005. Reconciling agricultural productivity and environmental integrity: a grand challenge for agriculture. *Front Ecol Environ* 3.1: 38–46.
- Rosa-Schleich J, Loos J, Mußhoff O, Tschamtké T. 2019. Ecological-economic trade-offs of diversified farming systems – A review. *Ecol Econ* 160: 251–263. <https://doi.org/10.1016/J.ECOLECON.2019.03.002>.
- Schiere JB, Lyklema J, Schakel J, Rickert KG. 1999. Evolution of farming systems and system philosophy. *Syst Res Behav Sci* 16.4: 375–390. [https://doi.org/10.1002/\(SICI\)1099-1743\(199907/08\)16:4%3C375::AID-SRES254%3E3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1099-1743(199907/08)16:4%3C375::AID-SRES254%3E3.0.CO;2-Q).
- Schillinger WF. 2011. Practical lessons for successful long-term cropping systems experiments. *Renew Agric Food Syst* 26: 1–3. <https://doi.org/10.1017/S1742170510000359>.
- Schwarzenbach RP, Egli T, Hofstetter TB, *et al.*, 2010. Global water pollution and human health. *Annu Rev Environ Resour* 35 : 109–136. <https://doi.org/10.1146/annurev-environ-100809-125342>.
- Sebillotte M. 1974. Agronomie et agriculture. Essai d'analyse des tâches de l'agronome. *Cahiers de l'ORSTOM* 24 : 3–25.
- Sebillotte M. 1990. Système de culture, un concept opératoire pour les agronomes. *Les systèmes de culture. Un Point sur... INRA*, 2-7380-0256-0.
- Siegel S, Castellan, Jr., NJ. 1988. *Nonparametric statistics for the behavioral sciences*, 1; 2nd ed. New York: McGraw-Hill.
- Silva V, Mol HGJ, Zomer P, Tienstra M., Ritsema CJ., Geissen V. 2019. Pesticide residues in European agricultural soils – A hidden reality unfolded. *Sci Total Environ* 653 : 1532–1545. <https://doi.org/10.1016/J.SCITOTENV.2018.10.441>.
- Smith RG, Warren ND, Cordeau S. 2020. Are cover crop mixtures better at suppressing weeds than cover crop monocultures? *Weed Sci* 68 : 186–194. <https://doi.org/10.1017/wsc.2020.12>.

- SSP - Agreste. 2019. Pratiques culturales en grandes cultures 2017 IFT et nombre de traitements. *Paris, France*. [En ligne] https://www.agreste.agriculture.gouv.fr/agreste-web/download/publication/publie/Chd1903/cd2019-3%20PK%20_%20janvier%202020%20v2.pdf
- Stanton RL, Morrissey CA, Clark RG. 2018. Analysis of trends and agricultural drivers of farmland bird declines in North America: A review. *Agric. Ecosyst. Environ* 254: 244–254.
- Tauvel P, Cadoux S, Duval R, *et al.*, Syppre: better reconciliation of global and local issues through innovative and diversified cropping systems. In Messéan A, Drexler D, Heim I, Paresys L, Stilmant D, Willer H, eds. European Conference on Crop Diversification. INRA and ÖMKI Budapest, Hungary, 2019, pp. 63–65.
- Tilman D, Balzer C, Hill J, Befort BL. 2011. Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci U S A* 108: 20260–20264. <https://doi.org/10.1073/PNAS.1116437108/-DCSUPPLEMENTAL>
- Tilman D, Cassman KG, Matson PA, Naylor R., Polasky S. 2002. Agricultural sustainability and intensive production practices. *Nature* 418: 671–677. <https://doi.org/10.1038/nature01014> 1.
- Toqué C, Cadoux S, Pierson P, *et al.*, SYPPRE: a project to promote innovations in arable crop production mobilizing farmers and stakeholders and including co-design, ex-ante evaluation and experimentation of multiservice farming systems matching with regional challenges. In: *5th International Symposium for Farming Systems Design. Agro2015, Montpellier, France*, 2015.
- Vasilachi IC, Tapciuc L, Chelaru GMF, *et al.*, Pesticides in the environment and harmonized risk indicators. In: *2020 International Conference on e-Health and Bioengineering (EHB) IEEE*, 2020, pp. 1–4.
- Vereijken, P. H. (1997). A methodical way of prototyping integrated and ecological arable farming systems (IEAFS) in interaction with pilot farms. *European Journal of Agronomy*, 7(1-3), 235–250
- Viguié L, Cavan N, Bockstaller C, *et al.*, 2021. Combining diversification practices to enhance the sustainability of conventional cropping systems. *Eur J Agron* 127: 126279. <https://doi.org/10.1016/j.eja.2021.126279>.
- Weber S, Jouy L, Angevin F, *et al.*, SYSTERRE[®], an online tool to describe diversified cropping systems, to calculate their performances, and assess their sustainability. In Messéan A, Drexler D, Heim I, Paresys L, Stilmant D, Willer H, eds. European Conference on Crop Diversification. INRA and ÖMKI Budapest, Hungary, 2019, pp. 337–338.
- Zampieri M, Weissteiner CJ, Grizzetti B, Toreti A., van den Berg F M., Dentener F. 2020. Estimating resilience of crop production systems: From theory to practice. *Sci Total Environ* 735: 139378. <https://doi.org/10.1016/j.scitotenv.2020.139378>.
- Zhang W, Ricketts TH, Kremen C, Carney K., Swinton S. M. 2007. Ecosystem services and dis-services to agriculture. *Ecol Econ* 64: 253–260. <https://doi.org/10.1016/j.ecolecon.2007.02.024>.
- Zuur AF, Ieno EN, Walker NJ, Saveliev, AA., Smith, GM. 2009. Mixed effects models and extensions in ecology with R. New York: 1; Springer.

Cite this article as: Longis S, Cadoux S, Toupet de Cordoue A-L, Tauvel P, Estienne M, Onzon P, Lescourret F, Rouillon C, Aubertot J-N, 2024. Performance of innovative cropping systems diversified with oilseeds and protein crops: identification and resolution of methodological issues, using the Syppre experimental network as a case study. *OCL* 31: 2.