

RESEARCH ARTICLE

Methodological considerations for redesigning sustainable cropping systems: the value of data-mining large and detailed farm data sets at the cropping system level

Nicolas MUNIER-JOLAIN (✉)¹, Martin LECHENET^{1,2}

¹ Agroecology Research Unit, AgroSup Dijon, INRAE. Univ. Bourgogne, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France

² Dijon Céréales, F-21600 Longvic, France

Abstract Redesigning cropping and farming systems to enhance their sustainability is mainly addressed in scientific studies using experimental and modeling approaches. Large data sets collected from real farms allow for the development of innovative methods to produce generic knowledge. Data mining methods allow for the diversity of systems to be considered holistically and can take into account the diversity of production contexts to produce site-specific results. Based on the very few known studies using such methods to analyze the crop management strategies affecting pesticide use and their effect on farm performance, we advocate further investment in the development of large data sets that can support future research programs on farming system design.

Keywords data mining, holistic, Integrated Pest Management, economics, DEPHY network

1 Introduction

Global agricultural production has increased over the last five decades, increasing the food availability per capita^[1]. However, agriculture worldwide is facing serious issues related to sustainability. The intensification of agriculture tends to increase the consumption of non-renewable resources, including energy (both for mechanization and for the synthesis and production of fertilizers, see Vlek et al.^[2], Deytieux et al.^[3]), phosphorus^[4] and fossil water^[5]. Agriculture contributes to the emission of greenhouse gases^[6], even though it could also contribute to the mitigation of carbon emissions through carbon sequestration^[7]. Fertilizer and pesticide intensification affects the quality of above- and below-ground water systems and

contributes to aquatic and terrestrial eutrophication^[8,9]. Pesticide residues contaminate the air with potential consequences for human health^[10]. Both the simplification of landscapes and the use of pesticides that affect non-target organisms might be responsible for severe losses in biodiversity in agricultural areas, hence reducing natural regulation of pests and increasing the risk of severe pest attacks^[11,12]. Moreover, pest control based on chemicals encourages the development of pest biotypes resistant to pesticide^[13], potentially resulting in severe threats for pest control over the medium to long-term.

The obvious lack of sustainability of most agricultural systems in industrialized countries calls for urgent improvement of both agriculture and the agrifood sector, in a way that will reconcile agricultural production with the protection of the environment, thereby enhancing the sustainability and long-term global food security. This is a complex challenge for the whole food system. Changes in agricultural systems must reduce their detrimental impacts while maintaining or improving profitability for farmers as well as the levels of agricultural production (both quantity and quality). Farming must continue to improve the provision of food for people—even though achieving food production and other requirements could also be reached through changes in human diets, see for example the prospective scenario of agrifood systems for Europe proposed by Poux and Aubert^[14].

The challenge of redesigning cropping systems to improve their sustainability in turn stimulates agricultural scientists to improve their methods for producing the knowledge needed by farmers to help them change. Farmers need documented results about which kind of combinations of technical management options, including technical innovations, would enhance overall sustainability without impairing profitability at the farm level. Of course, such results are likely to be very site-specific, depending on the soil and climate properties, and on the economic context of the farm. Policy makers are

Received July 15, 2019; accepted October 31, 2019

Correspondence: nicolas.munier-jolain@inra.fr

increasingly asking for more information about the consequences for the different components of sustainability, including the impacts on the environment, on non-renewable resources, and on biodiversity. The objective of this paper is to discuss the potential of methods based on data mining in large data sets compiling details of crop management at the farm level to produce valuable knowledge about systems sustainability. Such methods have a broad potential for development because large data sets at the cropping/farming system level are becoming available from participatory research programs involving farmers and farm advisers. We did not pretend to make an exhaustive literature review, but to discuss and advocate for a new way of producing knowledge in agronomy for helping farmers to change their practices, based on a few specific studies that used novel approaches. The point here is also not to describe the details of the methods, as the interested readers can find the information in the original articles, but rather to highlight how much such methods add to the range of investigation methods classically used by agronomists, that are based on factorial experiments, cropping system experiments and modeling. To the best of our knowledge, data mining has been applied to large data sets describing farmers' detailed practices at the cropping system level in only a few studies addressing the question of pesticide use^[15,16] and the relationship with farm productivity and profitability^[17]. The low number of such studies is probably related to the lack of large and detailed databases. In this paper, based on these rare studies, we discuss three critical assets of this scientific approach, namely, (1) results are directly connected with the reality of commercial farms, hence facilitating the use of the knowledge generated to support the transition at field level, (2) this method promotes a holistic view at the cropping/farming system level, which is a prerequisite for sound conclusions, and (3) the method makes it possible to explicitly take into account the diversity of production situations, i.e., the specific context of each farm site^[18], that might strongly impact the relationship between management strategy and the results of sustainability assessment. Prior to this, we discuss some strengths and weaknesses of experimental and modeling approaches regarding these three issues.

2 Factorial experiments, cropping system experiments and modeling

Factorial experiments have been and are still widely used by agronomists to investigate the effects of agricultural technologies and the interactions between technologies (e.g., to test cultivars at a range of sowing dates and fertilization levels^[19]). Experimental designs (e.g., randomized block design, split plot design, and others) and statistical methods (e.g., ANOVA) were developed specifically to analyze data produced by factorial experiments

and demonstrate the significance of observed results, i.e., not likely to be due to chance. The experimental approach is based on the *Ceteris paribus* principle of “all other things being equal”. The fact that results are shown not to be due to chance does not guarantee that results can be widely extrapolated, e.g., on a different sowing date or soil type, and thus would be the same on real commercial farms. The diversity of production situations is not considered unless the experiment is repeated over several years in a wide range of sites, which is often difficult to implement in practice. More importantly, factorial experiments can only manipulate a very limited number of factors at the same time, and therefore are not always useful in terms of promoting a holistic view of agricultural issues and to adapt all components of cropping systems in a consistent way, so as to improve the overall sustainability.

A wide range of long-term trials have been conducted all over the world during the past decades to evaluate the cumulative effects of crop rotations, fertilization regimes or soil tillage regimes on productivity, profitability, or other components of the agroecosystem^[20–23]. In most cases, long-term trials were based on factorial designs, which made it possible to demonstrate the effect of the factors tested, but such factorial designs obeying the rule of “all other things being equal” do not allow attention to be given to the consistency between the many components of the cropping system under consideration. Nonetheless, consistency of cropping systems (e.g., consistency between the sowing date, chosen cultivar and fertilization regime, consistency between crop rotation and soil tillage regime) is a prerequisite to optimize performance and progress toward sustainability^[24].

The issue of cropping system design according to sustainability targets has stimulated scientists to conduct long-term cropping-system experiments^[25–27]. In such experiments, the factors studied are not the individual technologies but the consistent cropping systems designed with complex combinations of technologies. This approach requires scrapping the rule of “all other things being equal” at the technology level, and therefore scrapping the objective of demonstrating the effects of individual technologies—a point that might be difficult to acknowledge for some scientists. Lechenet et al.^[24] made a comprehensive review of the diversity of methodologies and experimental layouts adopted for cropping system experiments. Two key points identified were (1) the definition of the concept of cropping system (either considered as a predefined sequence of techniques for a predefined crop rotation, or as a set of decision rules planning a flexibility of technologies applied for a predefined crop rotation, or as a set of decision rules planning both the flexibility of technologies and the flexibility of the crop sequence, as in many commercial farms), and (2) the holistic view of the consistency of components of the cropping systems. A valuable aspect of

the approach is that the data collected might be used for an assessment of different aspects of sustainability, including technical, economic and environmental performances^[3]. However, such heavy experimental arrangements are most often conducted in experimental stations with staff, equipment, organizational constraints and economic concerns quite different from real farms. Given the high requirements for space (for large replicated plots), time (for managing the decision rules) and human resources (for collecting and analyzing data), such cropping system experiments cannot be replicated at a large number of sites, raising the question of the generic value of the information produced.

Mechanistic cropping system models simulating the functioning of agroecosystems might be useful tools to explore the consequences of a wide range of strategic options based on various combinations of technical options and decision rules. Theoretically, this can be done for several production situations (soil types and climatic sequences^[28,29]). Models can help understanding the complex interactions within the agroecosystems, and therefore promote a holistic view of crop management. However, perfect simulation models will probably never be available for supporting cropping system design, as model developers have to find a compromise in the trade-off between (1) the number of processes simulated in the model, which is required for covering a large range of environmental situations and technical options, and (2) the risk of over-parameterization that would impair the predictive robustness of the model^[30].

3 Data mining

Considering the limitations of experimental and modeling approaches, data mining making use of large data sets collected from real cropping systems in real farms within participatory research projects could be a promising approach when addressing the issue of cropping system design for enhanced sustainability. Data mining covers a diversity of methods aiming at producing knowledge from large data sets without requiring any a priori hypotheses about data structure. The quality of the results therefore strongly depends on both the volume and the quality of the data set, and the robustness of the results has to be checked carefully to ensure a sound interpretation. Such methods have been used in agriculture wherever large data sets are available, for example to explore genomic data^[31], to make use of macro-economic data at the global level^[32], or of land use and spatial environmental information data^[33]. Data mining has been applied to precision agriculture and within field mapping based on remote sensing systems producing large amounts of data^[34]. However, data mining has rarely been used to address the sustainability of farmers' practices, considering the detailed technical components of cropping systems, because large data sets

that included the appropriate detailed information were lacking. We identified three rare exceptions of data mining focused on crop management, either at the crop level or at the cropping system level.

Colbach et al.^[35] used a large cropping system database mixing (1) real cropping systems described in real farms and (2) virtual cropping systems generated by random combinations of technical options. Weed control efficiency was evaluated for all these cropping systems using a simulation model, and decision trees were used to identify combinations of management techniques associated with multi-objective performance goals. Analyzing factors explaining the variability of pesticide use across a large number of winter wheat fields from commercial farms in North East Germany, Bürger et al.^[16] found that crop management (rotation, tillage, cultivar, seeding time and N fertilization) explained a lower part of pesticide use variance (about 10%) than farm characteristics and farmers' perception of risk. The statistical method was based on hierarchical variance partitioning by means of partial redundancy analysis.

With similar objectives but different data mining methods applied at the cropping system level, Lechenet et al.^[15,17] analyzed the data from the DEPHY network in France. This network is based on 3000 farmers who voluntarily redesigned their cropping systems with the explicit objective of reducing their reliance on pesticides. The network is organized to collect (1) details of crop management at the cropping system level, and (2) descriptors of the production situation for each farm. This produces a rare data set covering a large diversity of soils, climatic conditions, farming contexts, cropping systems and pesticide use, representing the existing diversity of farmer strategies to control pests, weeds, diseases, lodging and other constraints. The involvement of the farmers in the project was important because they sought to change their cropping system (crop rotation, crop management options), targeting reduced reliance on pesticides and enhanced sustainability. The diversity of their involvement and strategic choices contributed to the diversity within the produced database, which contributes to the power of the knowledge that the database can produce. Of course, the involvement of the farmers was also important during the data collection phase to produce the required large and detailed database. Lechenet et al.^[15] used a data mining statistical method, namely the classification and regression tree method, to identify various combinations of technical options associated with low vs. high pesticide use in a subset of around 1000 arable farms of the DEPHY network, from various production situations. They used random forest methods to rank the technical strategical options as a function of their impact on the reliance on pesticides. This ranking varied with the production situations, but most significant factors were (1) the crop diversification (with temporary grasslands whenever possible, with hardy and disease-resistant

crops and with crops with diversified sowing seasons), (2) the diversity of sown cultivars and the choice of cultivars with low susceptibility to diseases, (3) the delay in cereal sowing, (4) the reduction of pesticide doses, (5) the soil tillage regime, and (6) the nitrogen fertilization level^[15]. Therefore, the method confirmed the validity of most aspects of the Integrated Pest Management theory, but evidence was strengthened by the fact that it was based on real farms in real farming contexts. The weight of non-chemical pest management measures to explain pesticide use variance was much higher than in the previously cited German study^[16].

Using the same data set describing the diversity of strategic options at the cropping system level and the diversity of pesticide requirements and use, Lechenet et al.^[17] studied the correlation between pesticide reliance and two indicators of farm performance, namely the productivity and the profitability for the farmers. Given this relationship is likely to vary depending on the production situation (including soil type, climate and local markets), the statistical method used (i.e., the LASSO method^[36]) explicitly estimated the two-way interactions between variables describing the production situation and the pesticide use to explain farm performance. This method included a process to select the significant variables describing the production situation that explained between-site differences in the relationship between pesticide use and farm performance. The method was applied both at the cropping system and at specific crop (viz. wheat) levels.

In Fig. 1, each point corresponds to a single farm/site. The LASSO method allows the study, for each site, of the correlations between pesticide use and performance, for a sub-sample of farms sharing the same (or almost the same)

characteristics of context (soil, climate, socio-technical environment), with a higher weight in the regression for closer farms in the multidimensional neighborhood. Therefore, the method produces (surprisingly) one correlation value for each single site/farm. Green, yellow, red indicate negative, nil or positive correlation, respectively, between pesticide use and performance at the site, taking into account the specific context of the site (soil type, climate and economic environment). Level of pesticide use was estimated using a treatment frequency index. Productivity was computed as cumulative energy content in the harvested products at the cropping system level, or as yield in $t \cdot ha^{-1}$ for wheat. Profitability was estimated by the semi-net margin and expressed as $€ \cdot ha^{-1}$. The size of the points is proportional to the accuracy of the estimated correlations. More methodological details can be found in Lechenet et al.^[17].

At the cropping system level, results showed a limited proportion of sites with a positive correlation between level of pesticide use and either productivity (6% of sites, red dots in Fig. 1(a)) or profitability (22% of sites, red dots in Fig. 1(b)). In most situations, farmers using small amounts of pesticides had similar (yellow dots on Fig. 1) or greater (green dots) productivity or profitability than other farmers from the network in almost identical contexts^[17].

At the specific crop level, results were different. Most sites (73%) showed a positive correlation between pesticide use on wheat and wheat yield, because preventive measures adopted by farmers who wanted to reduce pest pressure (i.e., by the choice of cultivars driven by their resistance to diseases, and therefore not primarily driven by their yield potential, by delaying sowings and by the moderation of fertilization) also tended to decrease yield potential (Fig. 1(c)). However, the correlation between

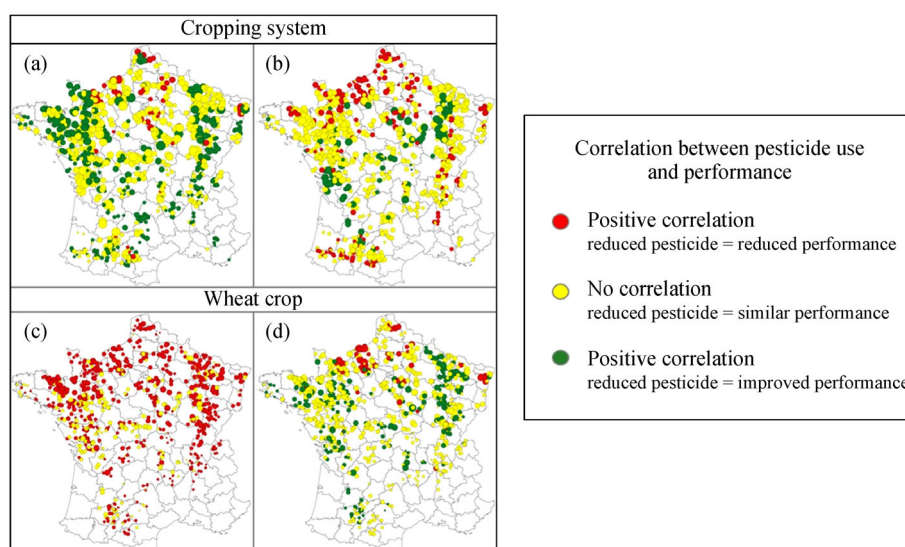


Fig. 1 Mapping of the correlations between pesticide use and productivity (a, c) and profitability (b, d), either computed at the cropping system level (a, b) or at the specific crop (viz. wheat) level (c, d). Adapted from Lechenet et al.^[17], with permission from Springer Nature.

pesticide use and profitability computed at the crop level was most often either nil (yellow dots in Fig. 1(d)) or negative (green dots), indicating that, in an average pricing context, reduced input costs offset lower yields in farms using small amounts of pesticide on wheat, and this contributed to maintain or even enhance profitability of the farm.

From this study, the results obtained at the cropping system level were previously published^[17], while the results at the crop (wheat) level are original previously unpublished ones. It is interesting to compare the results at both levels, because they are actually different, and this highlights how much it matters at what level economic analysis is done. At the crop level, decreases in crop yield associated with reduced pesticide use are offset by decreases in input costs (pesticide and fertilizers). At the cropping system levels, things are much more complicated, as decreases in pesticide use are associated with a number of differences, including different crop rotation, tillage regime, fertilization regimes, etc., which explains why the relationships between productivity and profitability are not simple and linear. In a few cases, higher productivity (expressed in harvested MJ·ha⁻¹) might be associated with lower gross proceeds and/or higher input costs (even if pesticide costs are lower) and/or higher mechanization costs, and therefore lower profitability. Based on the evidences that results depend on the scale, we advocate for prioritizing analysis of agricultural performances at the cropping system level (indeed more closely related with societal challenges) rather than at the single crop level.

This system-based approach to address the economic issues of pesticide use was novel. There is an abundant literature about pesticide use, pesticide overuse or pesticide underuse. However, most previous studies, primarily based on econometric approaches, estimated the marginal effects of pesticide use on either yield or profitability, most often only for a specific crop^[37–39]. The objective was to estimate the optimal level of pesticide in the given production context, implicitly considering “all other things, except pesticide, being equal” in the considered data set, whereas this assumption is typically not the reality across a large gradient of pesticide use in real farms. In contrast, the objectives of the study by Lechenet et al.^[17] was to investigate whether the in-depth redesigning of cropping systems for reducing pest pressure and pesticide requirements would impact economic performance at the farm level, which is critical information for any farmer seeking to enhance sustainability.

The question of pesticide use level has also been addressed in several studies, most often not considering the details of the cropping systems and their consistency. For example, Wu et al.^[40] analyzed the overuse of agricultural chemicals considering the farm size in a context of smallholder farms in China. In this specific context, the small size of farms is viewed as a limit impairing

knowledge transfer to farmers and the adoption of environmentally friendly practices.

4 Lessons learned

Methods based on data mining and on specific regression analysis allowing variable selection (the Lasso regression method) applied to large data sets collected from real farms within participatory research programs are powerful and complementary to other experimental methods to produce knowledge with generic value. They provide evidence that it is possible, in the context of western Europe, to adapt cropping systems so as to gain sustainability, with no negative impact on farm profitability. This allows for some factual elements in the current debate about the future of agriculture. The fact that the study was based on real farm data increases the power of the demonstration. The results also illustrate the key question of the scale at which sustainability issues are addressed. From the same set of farms, conclusions derived from analyses at cropping system level and at one specific crop level were different (i.e., no antagonism between reduced pesticide input and productivity at the farm level in most sites vs. a clear antagonism demonstrated at the wheat crop level). Scale matters when addressing the question of the relationship between technical options and the farm economics, and economic analysis should rather focus on the cropping/farming system complexity, whereas the question is often addressed at the crop level in the literature. The methods used explicitly considered cropping systems as complex combinations of technical components. They allowed the diversity of production situations to be taken into account to produce site-specific results.

Such methods are based on the value of large and detailed data sets describing the farmers’ practices at the cropping system level. We advocate the development of such large data sets, based on farm networks organized to collect data, using information systems designed for such purposes. There is a great need for standardized formats for agricultural data, in order to gather large scale data from a diversity of sources. Remote sensing and connected agricultural equipment could facilitate the development of such large databases in the near future, hence providing valuable data to produce more farm-based results and influence both farmers and policy makers to guide agriculture toward sustainability.

Acknowledgements We thank the farmers and farm advisors in the DEPHY network (Ministry of Agriculture, France), the SCEP-DEPHY research project funded by the AFB (French Agency for Biodiversity, France). Martin Lechenet’s PhD fellowship was funded by Agrosolutions (France). We thank Juliette Young and anonymous reviewers for comments and editings on the manuscript.

Compliance with ethics guidelines Nicolas Munier-Jolain and Martin Lechenet declare that they have no conflicts of interest or financial conflicts to

disclose.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

- Alexandratos N, Bruinsma J. World agriculture towards 2030/2050: the 2012 revision. Rome: FAO, 2012
- Vlek P L G, Rodriguez-Kuhl G, Sommer R. Energy use and CO₂ production in tropical agriculture and means and strategies for reduction or mitigation. *Environment, Development and Sustainability*, 2004, **6**(1–2): 213–233
- Deytieux V, Nemecek T, Freiermuth Knuchel R, Gaillard G, Munier-Jolain N. Is Integrated Weed Management efficient for reducing environmental impacts of cropping systems? A case study based on life cycle assessment. *European Journal of Agronomy*, 2012, **36**(1): 55–65
- Schröder J J, Smit A L, Cordell D, Rosemarin A. Improved phosphorus use efficiency in agriculture: a key requirement for its sustainable use. *Chemosphere*, 2011, **84**(6): 822–831
- Pimentel D, Houser J, Preiss E, White O, Fang H, Mesnick L, Barsky T, Tariche S, Schreck J, Alpert S. Water resources: agriculture, the environment, and society. *Bioscience*, 1997, **47**(2): 97–106
- Robertson G P, Paul E A, Harwood R R. Greenhouse gases in intensive agriculture: contributions of individual gases to the radiative forcing of the atmosphere. *Science*, 2000, **289**(5486): 1922–1925
- Lal R. Soil carbon sequestration impacts on global climate change and food security. *Science*, 2004, **304**(5677): 1623–1627
- Silva E, Daam M A, Cerejeira M J. Aquatic risk assessment of priority and other river basin specific pesticides in surface waters of Mediterranean river basins. *Chemosphere*, 2015, **135**: 394–402
- Wang J, Fu Z, Qiao H, Liu F. Assessment of eutrophication and water quality in the estuarine area of Lake Wuli, Lake Taihu, China. *Science of the Total Environment*, 2019, **650**(Pt 1): 1392–1402
- Hamsan H, Ho Y B, Zaidon S Z, Hashim Z, Saari N, Karami A. Occurrence of commonly used pesticides in personal air samples and their associated health risk among paddy farmers. *Science of the Total Environment*, 2017, **603–604**: 381–389
- Bianchi F J J A, Booij C J H, Tschamtké T. Sustainable pest regulation in agricultural landscapes: a review on landscape composition, biodiversity and natural pest control. *Proceedings. Biological Sciences*, 2006, **273**(1595): 1715–1727
- Martel G, Aviron S, Joannon A, Lalechère E, Roche B, Boussard H. Impact of farming systems on agricultural landscapes and biodiversity: from plot to farm and landscape scales. *European Journal of Agronomy*, 2019, **107**: 53–62
- Gould F, Brown Z S, Kuzma J. Wicked evolution: can we address the sociobiological dilemma of pesticide resistance? *Science*, 2018, **360**(6390): 728–732
- Poux X, Aubert P M. An agroecological Europe in 2050: multifunctional agriculture for healthy eating—findings from the Ten Years For Agroecology (TYFA) modelling exercise. Paris: *IDDRI*, 2018
- Lechenet M, Makowski D, Py G, Munier-Jolain N. Profiling farming management strategies with contrasting pesticide use in France. *Agricultural Systems*, 2016, **149**: 40–53
- Bürger J, de Mol F, Gerowitt B. Influence of cropping system factors on pesticide use intensity—a multivariate analysis of on-farm data in North East Germany. *European Journal of Agronomy*, 2012, **40**: 54–63
- Lechenet M, Dessaint F, Py G, Makowski D, Munier-Jolain N. Reducing pesticide use while preserving crop productivity and profitability on arable farms. *Nature Plants*, 2017, **3**: 17008,
- Aubertot J N, Robin M H. 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*, 2013, **8**(9): e73202
- Lu D, Lu F, Pan J, Cui Z, Zou C, Chen X, He M, Wang Z. The effects of cultivar and nitrogen management on wheat yield and nitrogen use efficiency in the North China Plain. *Field Crops Research*, 2015, **171**: 157–164
- Soane B D, Ball B C. Review of management and conduct of long-term tillage studies with special reference to a 25-yr experiment on barley in Scotland. *Soil & Tillage Research*, 1998, **45**(1–2): 17–37
- Richter D D B, Hofmockel M, Callahan M A, Powlson D S, Smith P. Long-term soil experiments: keys to managing Earth’s rapidly changing ecosystems. *Soil Science Society of America Journal*, 2007, **71**(2): 266–279
- Edward Johnston A. The value of long-term field experiments in agricultural, ecological, and environmental research. *Advances in Agronomy*, 1997, **59**: 291–333
- Varvel G E. Rotation and nitrogen fertilization effects on changes in soil carbon and nitrogen. *Agronomy Journal*, 1994, **86**(2): 319–325
- Lechenet M, Deytieux V, Antichi D, Aubertot J N, Bàrberi P, Bertrand M, Cellier V, Charles R, Colnenne-David C, Dachbrodt-Saaydeh S, Debaeke P, Doré T, Farcy P, Fernandez-Quintanilla C, Grandeau G, Hawes C, Jouy L, Justes E, Kierzek R, Kudsk P, Lamichhane J R, Lescourret F, Mazzoncini M, Melander B, Messéan A, Moonen A C, Newton A C, Nolot J M, Panozzo S, Retaureau P, Sattin M, Schwarz J, Toqué C, Vasileiadis V P, Munier-Jolain N. Diversity of methodologies to experiment Integrated Pest Management in arable cropping systems: analysis and reflections based on a European network. *European Journal of Agronomy*, 2017, **83**: 86–99
- Teasdale J R, Coffman C B, Mangum R W. Potential long-term benefits of no-tillage and organic cropping systems for grain production and soil improvement. *Agronomy Journal*, 2007, **99**(5): 1297–1305
- Debaeke P, Munier-Jolain N, Bertrand M, Guichard L, Nolot J M, Faloya V, Saulas P. Iterative design and evaluation of rule-based cropping systems: methodology and case studies—a review. *Sustainable Agriculture*, 2009, **29**(1): 707–724
- Chikowo R, Faloya V, Petit S, Munier-Jolain N. Integrated Weed Management systems allow reduced reliance on herbicides and long-term weed control. *Agriculture, Ecosystems & Environment*, 2009, **132**(3–4): 237–242
- Zhang Y, Feng L P, Wang J, Wang E L, Xu Y L. Using APSIM to explore wheat yield response to climate change in the North China Plain: the predicted adaptation of wheat cultivar types to

- vernalization. *Journal of Agricultural Science*, 2013, **151**(6): 836–848
29. Colbach N, Biju-Duval L, Gardarin A, Granger S, Guyot S H M, Mézière D, Munier-Jolain N, Petit S. The role of models for multicriteria evaluation and multiobjective design of cropping systems for managing weeds. *Weed Research*, 2014, **54**(6): 541–555
 30. Colbach N. Modelling cropping system effects on crop pest dynamics: how to compromise between process analysis and decision aid. *Plant Science*, 2010, **179**(1–2): 1–13
 31. Naika M, Shameer K, Sowdhamini R. Comparative analyses of stress-responsive genes in *Arabidopsis thaliana*: insight from genomic data mining, functional enrichment, pathway analysis and phenomics. *Molecular BioSystems*, 2013, **9**(7): 1888–1908
 32. Shortridge J E, Falconi S M, Zaitchik B F, Guikema S D. Climate, agriculture, and hunger: statistical prediction of undernourishment using nonlinear regression and data-mining techniques. *Journal of Applied Statistics*, 2015, **42**(11): 2367–2390
 33. Tayyebi A, Pijanowski B C, Linderman M, Gratton C. Comparing three global parametric and local non-parametric models to simulate land use change in diverse areas of the world. *Environmental Modelling & Software*, 2014, **59**: 202–221
 34. Waheed T, Bonnell R B, Prasher S O, Paulet E. Measuring performance in precision agriculture: CART-A decision tree approach. *Agricultural Water Management*, 2006, **84**(1–2): 173–185
 35. Colbach N, Colas F, Pointurier O, Queyrel W, Villerd J. A methodology for multi-objective cropping system design based on simulations. Application to weed management. *European Journal of Agronomy*, 2017, **87**: 59–73
 36. Tibshirani R. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B: Methodological*, 1996, **58**(1): 267–288
 37. Lichtenberg E, Zilberman D. The econometrics of damage control: why specification matters. *American Journal of Agricultural Economics*, 1986, **68**(2): 261–273
 38. Fernandez-Cornejo J, Jans S, Smith M. Issues in the economics of pesticide use in agriculture: a review of the empirical evidence. *Applied Economic Perspectives and Policy*, 1998, **20**(2): 462–488
 39. Wang J, Chu M, Ma Y. Measuring rice farmer’s pesticide overuse practice and the determinants: a statistical analysis based on data collected in Jiangsu and Anhui Provinces of China. *Sustainability*, 2018, **10**(3): 677
 40. Wu Y, Xi X, Tang X, Luo D, Gu B, Lam S K, Vitousek P M, Chen D. Policy distortions, farm size, and the overuse of agricultural chemicals in China. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, **115**(27): 7010–7015