Institut für Lebensmittel- und Ressourcenökonomik (ILR)

Advances in technology evaluation and decision support studies using bio-economic farm models

D i s s e r t a t i o n
zur
Erlangung des Grades
Doktor der Agrarwissenschaften
(Dr.agr.)

der
Landwirtschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von

Christoph Pahmeyer

aus

Halle (Westf.)

Bonn, 2022

Referent: PD Dr. Wolfgang Britz

Korreferent: Prof. Dr. Thomas Heckelei

Tag der mündlichen Prüfung: 25.03.2022

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät der Universität Bonn

Danksagung

An dieser Stelle möchte ich mich bei meinen Kollegen-/innen, Verwandten und Freunden-/innen bedanken, die mich bei der Fertigstellung dieser Dissertation tatkräftig unterstützt haben. Für die Betreuung meiner Dissertation, Projekttreffen kombiniert mit Campingausflügen, die Freiheiten bei der Themensetzung sowie das entgegengebrachte Vertrauen möchte ich mich herzlich bei PD Dr. Wolfgang Britz bedanken. Herrn Prof. Dr. Thomas Heckelei möchte ich für das entgegengebrachte Vertrauen im PhenoRob Projekt sowie die Übernahme des Korreferats danken. Darüber hinaus danke ich meinen Kollegen-/innen vom Institut für Lebensmittel- und Ressourcenökonomik (ILR) für vier wunderbare Jahre, spannende Diskussionen und After-Work-Events. Ein großes Dankeschön geht sowohl an die Mensa-Gang für die entspannenden und sättigenden Mittagspausen, als auch an das gesamte Mensapersonal, das mich über Jahre mit Nudeln, Pesto und ACE-Saft versorgt hat.

Ein besonderer Dank gilt Till Kuhn, ohne dessen Ratschläge, Ideen und Kreativität die Dissertation vermutlich erst deutlich später fertigstellt worden wäre. Auch möchte ich mich bei meinen zukünftigen Co-Autoren/-innen Lin Mei Chang, Sebastian Rasch und Hugo Storm für die vielen spannenden inhaltlichen Diskussionen sowie das thematische Feedback bedanken. Meiner Bürokollegin Julia Heinrichs danke ich für den vielen Spaß im Büro, Feedback auf dem kurzen Dienstweg und Unmengen an Schokolade. Lennart Kokemohr danke ich insbesondere für die vieljährige Unterstützung im SustainBeef-Projekt sowie die schönen Stunden außerhalb des Büros auf dem Rad und beim Kochen. Bei Rienne Wilts möchte ich mich für das unermüdliche Korrekturlesen, ihr offenes Gehör sowie die ein oder andere Kuchenpause bedanken.

Nicht zuletzt möchte ich mich für die jahrelange Unterstützung bei meiner Familie bedanken. Bei meinen Geschwistern Christine, Marlene und Simon und ganz besonders bei meinen Eltern dafür, dass sie mir den Spaß und das Interesse an der Landwirtschaft vermittelt, meine Begeisterung gefördert und erhalten haben und mir immer wieder gezeigt haben, worauf es wirklich ankommt.

Abschließend möchte ich mich bei meiner Freundin Sophia für die Unterstützung während meiner Promotionszeit, ihrer Toleranz bezüglich meines schlechten Zeitmanagements, aber vor allem für ihren aufmunternden Charakter und die vielen schönen Ablenkungen von der Arbeit bedanken.

Abstract

The use of so-called bio-economic farm models makes it possible to determine the most cost-effective adaptation strategy to novel policy measures and to evaluate whether the use of new technologies can be economically viable for a farm. This thesis aims to show potential ways for modeling novel policies and technologies in the existing farm model FarmDyn as well as the decision support system ‘Fruchtfolge’ developed in the context of the thesis.

As an example of a novel technology with far-reaching consequences on a farm's production process, the thesis highlights the economical optimal use of sex-sorted semen and crossbreeding among a dairy farm population from the German federal state of North Rhine-Westphalia. Using the holistic farm model FarmDyn, potential profit increases ranging from 0 €/cow/year to 568 €/cow/year, with an average of 79 €/cow/year are shown. The results demonstrate that modern breeding technologies have the potential to improve dairy farm profits, although they must be viewed in the context of farm-specific production settings.

To assess a policy with measures targeting single fields of a farm in a bio-economic farm model, the thesis presents a novel decision support system called ‘Fruchtfolge’. The model assists farmers with finding a cost-minimal adoption strategy to the newly revised Fertilization Ordinance in Germany. The decision support system presents farmers a cropping choice and fertilization management recommendation for each of their fields. In a case study application involving a farm managing fields both outside and inside of a nitrate sensitive area, the DSS is shown to mitigate the farms' compliance costs to the revised Fertilization Ordinance by more than 5% when compared to the former optimal resource allocation.

A major difference between the models FarmDyn and ‘Fruchtfolge’ lies in the spatial resolution of the cropping decision. To quantify the influence of the spatial resolution on the simulation results, high-resolution farm data are required. In the context of the thesis, a methodology for generating such a dataset for the German federal state of North Rhine-Westphalia is presented and used to analyse the aforementioned influence on the simulation results of bio-economic models. The findings indicate that crop choices per farm differ by 11% on average, resulting in profit differences ranging from -306 €/ha to 434 €/ha when explicitly modeling single plots of a farm, compared to a traditional approach where crop shares on all arable land are modeled.

The bio-economic models presented in this thesis can contribute to both ex-ante and ex-post analysis of the economic impact of novel policies and technologies on farms. They can also be applied to other research questions and may help to disseminate knowledge to farmers and farm advisers.

Kurzfassung

Der Einsatz sogenannter bioökonomischer Betriebsmodelle ermöglicht es, die kosteneffizienteste Anpassungsstrategie an neue Politikmaßnahmen zu bestimmen und zu evaluieren, ob die Nutzung neuer Technologien für einen landwirtschaftlichen Betrieb ökonomisch sinnvoll sein kann. Das Ziel der vorliegenden Arbeit ist es, Fallbeispiele für Politik- und Technologieevaluierungen mithilfe des bestehenden Betriebsmodells FarmDyn sowie des im Rahmen der Arbeit neu entwickelte Entscheidungshilfesystems „Fruchtfolge“ aufzuzeigen.

Als ein Beispiel für eine Technologieevaluierung wird der ökonomisch optimale Einsatz von sogenanntem gesextem Sperma sowie von Kreuzzüchtungen in Nordrhein-Westfälischen Milchviehbetrieben beleuchtet. Anhand der Anwendung des Betriebmodells FarmDyn zeigen sich potenzielle Gewinnsteigerungen von 0 €/Kuh/Jahr bis 568 €/Kuh/Jahr bei einem durchschnittlichen Gewinnzuwachs von 79 €/Kuh/Jahr. Die Ergebnisse verdeutlichen, dass moderne Züchtungstechnologien das Potenzial haben, Profite in der Milchviehwirtschaft zu steigern, jedoch stets im speziellen Kontext einzelner Betriebe evaluiert werden müssen.

Um eine Politikmaßnahme mit speziellen Reglementierungen auf Schlagebene abzubilden, wird das Entscheidungshilfesystem „Fruchtfolge“ vorgestellt, welches Landwirt(e)-/innen dabei unterstützt, die kostengünstigste Anpassungsstrategie an die letzte Revision der Düngeverordnung zu finden. Dabei erstellt das System eine Anbau- und Düngeempfehlung für jede Fläche des Betriebs. Für einen Fallstudienbetrieb mit Flächen sowohl innerhalb als auch außerhalb eines nitratbelasteten „roten Gebietes“ kann gezeigt werden, dass der aus der Anwendung resultierende Anbauplan ca. 5% geringere Anpassungskosten aufweist als der vorherige Anbauplan des Betriebes.

Die Betriebsmodelle FarmDyn und „Fruchtfolge“ unterscheiden sich vornehmlich in Bezug auf die räumliche Auflösung der Anbauplanungsentscheidung. Um den Einfluss der räumlichen Auflösung auf die Simulationsergebnisse zu quantifizieren, werden hochaufgelöste Betriebsdaten benötigt. Im Rahmen der Arbeit wird eine Methodik zur Generierung eines solchen Datensatzes beispielhaft für Nordrhein-Westfalen vorgestellt und dieser verwendet, um den zuvor genannten Einfluss in bioökonomischen Modellen zu analysieren. Die Ergebnisse verdeutlichen, dass die Kulturanteile zwischen einem Ansatz auf Schlagebene verglichen mit einem aggregierten Ansatz im Durchschnitt um 11% abweichen, mit einhergehenden Profitunterschieden zwischen -306 €/ha bis 434 €/ha.

Die im Rahmen der Arbeit vorgestellten bioökonomischen Modelle können sowohl zur Ex-ante- als auch zur Ex-post-Analyse der ökonomischen Auswirkungen von neuen Politikmaßnahmen und Technologien auf landwirtschaftliche Betriebe beitragen. Ebenfalls können sie auf andere Forschungsfragen übertragen werden und helfen, die gewonnenen Erkenntnisse an Landwirt(e)-/innen weiterzugeben.

Contents

List of tables

List of figures

Abbreviations

ABM	Agent based model
BEFM	Bio-economic farm model
CER	Cumulative energy requirement
CRN	Customer reference number
DSS	Decision support system
FO	Fertilization Ordinance
FSS	Farm Structure Survey
GUI	Graphical User Interface
HPD	Highest posterior density
IACS	Integrated administration and control system
IE	Insemination effort
LAU	Local administrative unit
LHS	Latin hypercube sampling
LP	Linear programming
MILP	Mixed integer linear programming
N	Nitrogen
N_{min}	Mineralized nitrogen
NRW	North Rhine-Westphalia
NPV	Net present value
NUTS	Nomenclature of territorial units for statistics
P	Phosphate
PMP	Positive mathematical programming
S.D.	Standard deviation
SQR	Soil quality rating

Chapter 1
Introduction

Farming systems are known to be a major driver of environmental degradation (Springmann et al., 2018). As one of the results of environmental degradation, the total utilized agricultural area is shrinking throughout the world (Bruinsma, 2017). Despite this circumstance, a growing world population needs to be fed which is challenging the agricultural sector to simultaneously become more sustainable and resilient, as well as more productive (Bruinsma, 2017). In order to reach these goals, policies regimenting farming practices associated with negative externalities, and fostering practices enabling higher productivity and sustainability are implemented throughout the world (Troost et al., 2015). In the European Union (EU) for instance, the Green Deal and farm-to-fork policy initiatives, as well as the renewed Common Agricultural Policy (CAP) (2023-2027) present targeted emission reduction goals, as well as desired farming practices for the years to come. On the federal level, member states have to implement the EU Nitrates Directive aiming to protect the water quality across Europe (European Commission, 2010).

Furthermore, novel technologies and innovations targeting an increase in productivity while reducing the environmental burden enter the agricultural market (Sparrow and Howard, 2021). Such technologies and innovations cover a wide range of possible applications in the agricultural domain, spanning from innovative breeding methods over autonomous field equipment to digital tools such as decision support systems (DSS) (Lowenberg-DeBoer et al., 2020; Rose et al., 2016; Vishwanath and Moreno, 2018).

Especially from an economic point of view, these novel policies and technologies pose substantial challenges for farmers (Beckman et al., 2021). Finding a cost-minimizing compliance strategy to a policy or deciding whether adopting a new technology or farming practice is viable can be a challenging task, which has to be evaluated in the specific context of an individual farm (Chavas and Nauges, 2020; Kuhn et al., 2019). A further aspect that complicates the evaluation is the (spatial) resolution of the aforementioned policies and technologies: More often, they target specific production processes or (sub-) regions, thus affecting each farm differently. The effects vary depending on the contribution of the production process to the overall farm profitability or the share of the farms plots falling into a region targeted by the policy. Given the complexity and far reaching impacts these policies and technologies can have on a farm, assessing their economic impact is not only crucial to farmers. Policy makers seeking to alleviate economic losses induced by environmental policies or interested in fostering novel potentially sustainable farming technologies can also benefit from these assessments (Janssen and van Ittersum, 2007).

From a technical point of view, assessing the potential economic impact of a policy or a technology on a farm can be challenging. Given the interdependences of processes and resources, conceptual models are required in order to reduce the overall complexity to a comprehensible level (Sørensen et al., 2010). In the context of policy and technology evaluation, especially bio-economic farm models (BEFM) are frequently employed in the scientific literature. A BEFM is "defined as a [mathematical] model that links formulations describing farmers' resource management decisions to formulations that describe current and alternative production possibilities in terms of required inputs to achieve certain outputs and associated externalities" (Janssen and van Ittersum, 2007, p. 623). In addition to being used in studies concerning the evaluation of policies or technologies, BEFMs often form the basis for DSS for farmers and farm advisors (Jones et al., 2017). Such models can show potential profit maximizing adaptation strategies to a policy, or whether a technology could be economically profitable for a specific farm or not (Antle et al., 2017).

1.1 Motivation

Given their fine-grained resolution, the aforementioned novel policies and technologies challenge the way BEFMs are developed and used in their current generation (Britz et al., 2021). For a thorough economic analysis, the underlying bio-economic relationships need to be incorporated in high detail, while still reflecting a holistic image of the farm. Balancing these conditions in policy and technology evaluation studies, as well as DSS, poses new challenges regarding data-, resource-, and technology requirements. In the context of these studies, extensions to existing BEFMs are required that integrate the policies and technologies in the needed level of detail, but also allow for an evaluation of these on a whole farm population scale in order to capture farm heterogeneity. In the context of DSS, where BEFMs have the potential to mitigate potential compliance costs to novel policies, they are required to improve the communication of results to their users in order to increase adoption levels (Jones et al., 2017).

A further issue concerning both DSS and BEFMs used in technology and policy adoption studies is a lack of adequate and detailed data (Jones et al., 2017). This lack concerns both the availability of data considering farms and their endowments, as well as general farm planning data required to parameterise the BEFMs (Reidsma et al., 2018). Considering the required level of detail of many novel policies and technologies, such data is needed in order to analyse the impact of these policies and technologies on a whole farm population. However, with the increasing availability of public datasets that encompass agriculturally relevant data, new possibilities for overcoming the issues caused by data scarcity emerge. One the one hand, linked-open-datasets concerning farm planning data can be utilised in order to improve the depiction of technical details and varying farm characteristics in BEFMs (Martini et al., 2014). On the other hand, existing farm typologies (e.g. Kuhn and Schäfer (2018)) can be extended with spatial data concerning i.a. plot geometries and soil qualities in order to attain spatially explicit farm populations that allow for a more detailed analysis of the impact of the aforementioned novel policies and technologies.

With better data availability, possibilities for depicting certain on-farm decision processes with higher accuracy emerge. An example for a central on-farm decision process that is depicted in both BEFMs used for policy and technology evaluation, as well as in DSS, is the choice of crops planted on a farms fields. Traditionally, BEFMs depict the choice of crops using crop shares of a farms aggregate land endowment, whereas highly detailed farm data allow for a depiction of the decision problem on a single plot level. However, with an increasing level of detail, data- and resource requirements of a BEFM may also drastically increase. Furthermore, result interpretation can become more complex, as additional variables and interdependencies are entered into the model. Depending on the use-case and research question, this additional level of detail may therefore not always lead to substantial additional insights. Consequently, there is a need for a study clarifying which type of analysis or use-case can benefit from the higher level of simulation detail, and for which scenarios the higher level of detail may not be required.

1.2 Research aims

The dissertation aims to show potential ways for modeling highly detailed policies and technologies in BEFMs used for policy and technology analysis as well as for decision support. As the methodological focus of the dissertation primarily lies in the development, extension, and application of BEFMs, use-cases for both existing and novel BEFMs are explored.

For the further development of an existing BEFM, the model FarmDyn (Britz et al., 2018) is used in the dissertation. The BEFM FarmDyn is a highly detailed single farm model. It has originally been developed for assessing marginal abatement costs of greenhouse gases on dairy farms (Lengers et al., 2014). Since its inception, multiple extensions to the model have been made, introducing support for new farm branches (Garbert, 2013; Schäfer et al., 2017), fertilizer- and environmental policies (Kuhn et al., 2019), as well as optional integrations with other agricultural systems models (Kuhn et al., 2020). FarmDyn can be used to model a wide range of agricultural systems out-of-the-box, while extensions to the model can be integrated due to its template based design (Britz et al., 2018). As proposed by Reidsma et al. (2018), using an existing and established BEFM for policy and technology evaluation studies should be encouraged over creating a new model, as model development time can be augmented on the given research problem at hand, and simulation results can be validated more easily.

By design, the FarmDyn BEFM is primarily aimed to be operated by researchers. With its multiple methodological extensions and included farm branches, the model is well suited for policy and technology evaluation studies, but less fitting as a basis for a DSS. Since DSS are generally aimed to be operated by farmers and farm advisers, and specifically target a certain decision problem, the multi-purpose BEFM FarmDyn can be considered a less suitable option with regards to DSS. In the context of the thesis, a novel DSS called ‘Fruchtfolge’ is established. ‘Fruchtfolge’ is a web-based DSS specifically focusing on the crop choice and accompanying fertilization decision problem. Opposing the methodology used in FarmDyn, ‘Fruchtfolge’ depicts the crop choice problem and related fertilization management on the individual plot level. Farm endowments and planning data are automatically queried from various linked open databases. However, ‘Fruchtfolge’ does not explicitly incorporate animal husbandry, investments, or other farm branches found in holistic farm models such as FarmDyn.

In order to fulfil the research aims, the thesis gives practical examples for implementing BEFMs given the various use-cases: As an example for a novel technology with far reaching consequences on a farms production process, the dissertation highlights how an advanced cattle breeding technology can be incorporated in a BEFM considering its complex interactions on bio-physical as well as economic processes of a farm. As a specific example of such a technology in the dairy sector, the dissertation focuses on the inclusion of so-called "sexed semen”, as well as crossbreeding as two novel cattle breeding strategies in the existing model FarmDyn. Therefore, the aim of the first article presented in Chapter 2 can be summarized as:

I. *Implement novel breeding technologies and strategies with highly detailed underlying biological relationships in FarmDyn as a holistic farm model.*

A further goal of the dissertation is to show a potential way for modeling a novel policy in a BEFM. In this regard, the dissertation shows the novel DSS 'Fruchtfolge' that helps farmers to find a cost minimal adoption strategy to the revised version of the German Fertilizer Ordinance (FO) with minimal effort. As a policy, the German FO (revision of 2020) is chosen, as it incorporates both measures targeting the farm as whole, as well as detailed measures targeting individual plots of a farm (BMEL, 2020). Considering this policy, the question of which crops to grow on which plot, as well as how these crops should be fertilized becomes more complex for decision makers. As finding a cost-minimal abatement strategy to the policy is far from trivial, a DSS can help reduce the additional burden on decision makers, and potentially also lower compliance costs. In the context of the DSS, especially the communication of the simulation results to end users, as well as a simplified interface for including the endowments of a specific farm are targeted as a central research point. In order to achieve this goal, the web-based DSS 'Fruchtfolge' is established, incorporating the depiction of crop choices on the single plot level, as well as automated data acquisition and a user-centered design (Rose et al., 2016). Therefore, the research aim of the second article presented in Chapter 3 is:

II. *Develop a DSS for cropping choices based on big data and user-centered design, and illustrate the benefits of the DSS in a case study application on the German FO.*

Policies such as the revised version of the German FO do not only introduce measures targeting farms as a whole, but also measures targeting single plots of a farm. Given the overall scarcity of highly detailed single farm data mentioned earlier, this additional level of detail proves it difficult to assess the impacts of such a policy on a whole farm population. In order to overcome this data limitation, the dissertation aims to provide a methodology as well as an accompanying dataset merging an existing farm typology (Kuhn and Schäfer, 2018) with spatially explicit open datasets in order to achieve a synthetic farm population. For those reasons, the research aim of the third article presented in Chapter 4 is:

III. *Provide data and a methodology for creating a synthetic farm population with single farm data of the German federal state of North Rhine-Westphalia (NRW).*

A major motivation for the inception of the 'Fruchtfolge' DSS is the detailed depiction of the cropping choice and related fertilization management problem on the single plot level. Opposed to this approach, holistic farm BEFMs such as FarmDyn frequently involve a certain level of land aggregation, as spatially explicit data for modeling the crop choice problem on the individual field level are rarely available. Given the highly detailed synthetic farm population, new possibilities for the analysis of novel policies and technologies emerge. From a methodological point of view, many of these policies and technologies directly or indirectly affect the choice of crops in a BEFM. The effect of these different levels of land aggregation on the simulation results has not been studied yet. Therefore, the research aim of the fourth article presented in Chapter 5 is:

IV. *Evaluate the current approaches to model crop choices in BEFMs, and assess how these approaches affect BEFM simulation results in both technology and policy evaluation studies as well as DSS.*

1.3 **Proceedings**

The dissertation is structured as follows: Chapter 2 implements and assesses the impact of the novel breeding technology "sexed semen", as well as the breeding strategy crossbreeding on dairy farms in the German federal state of North Rhine-Westphalia. Chapter 3 presents the novel DSS "Fruchtfolge", and evaluates its potential to mitigate compliance costs for a case-study farm under different scenarios in the German federal state of North Rhine-Westphalia. Chapter 4 gives a methodology for creating a synthetic farm population using a farm typology and other spatially explicit datasets. Chapter 5 analyses the difference in BEFM simulation results between different levels of detail with regards to crop choices, and gives recommendations for which option to use for which research question. Eventually, Chapter 6 concludes.

1.4 References

Antle, J.M., Basso, B., Conant, R.T., Godfray, H.C.J., Jones, J.W., Herrero, M., Howitt, R.E., Keating, B.A., Munoz-Carpena, R., Rosenzweig, C., Tittonell, P., Wheeler, T.R., 2017. Towards a new generation of agricultural system data, models and knowledge products: Design and improvement. Agricultural Systems 155, 255–268. https://doi.org/10.1016/j.agsy.2016.10.002

Beckman, J., Ivanic, M., Jelliffe, J., 2021. Market impacts of Farm to Fork: Reducing agricultural input usage. Applied Economic Perspectives and Policy n/a. https://doi.org/10.1002/aepp.13176

BMEL, 2020. Verordnung zur Änderung der Düngeverordnung und anderer Vorschriften. Federal Ministry of Food and Agriculture (BMEL). [WWW Document]. URL https://www.bundesrat.de/SharedDocs/drucksachen/2020/0001-0100/98-20.pdf?__blob=publicationFile&v=2 (accessed 6.5.20).

Britz, W., Ciaian, P., Gocht, A., Kanellopoulos, A., Kremmydas, D., Müller, M., Petsakos, A., Reidsma, P., 2021. A design for a generic and modular bio-economic farm model. Agricultural Systems 191, 103133. https://doi.org/10.1016/j.agsy.2021.103133

Britz, W., Lengers, B., Kuhn, T., Schäfer, D., 2018. A highly detailed template model for dynamic optimization of farms. Institute for Food and Resource Economics, University of Bonn. Model Documentation.

Bruinsma, J. (Ed.), 2017. World Agriculture: Towards 2015/2030: An FAO Study. Routledge, London. https://doi.org/10.4324/9781315083858

Chavas, J.-P., Nauges, C., 2020. Uncertainty, Learning, and Technology Adoption in Agriculture. Applied Economic Perspectives and Policy 42, 42–53. https://doi.org/10.1002/aepp.13003

European Commission, 2010. The EU Nitrates Directive [WWW Document]. URL https://ec.europa.eu/environment/pubs/pdf/factsheets/nitrates.pdf (accessed 10.11.21).

Garbert, J., 2013. Ökonomische Auswirkungen von Politiken zur Umsetzung der Wasserrahmenrichtlinie auf die Schweinehaltung im Münsterland.

Garske, B., Bau, A., Ekardt, F., 2021. Digitalization and AI in European Agriculture: A Strategy for Achieving Climate and Biodiversity Targets? Sustainability 13, 4652. https://doi.org/10.3390/su13094652

Janssen, S., van Ittersum, M.K., 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems 94, 622–636. https://doi.org/10.1016/j.agsy.2007.03.001

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017. Brief history of agricultural systems modeling. Agricultural Systems 155, 240–254. https://doi.org/10.1016/j.agsy.2016.05.014

Kuhn, T., Enders, A., Gaiser, T., Schäfer, D., Srivastava, A.K., Britz, W., 2020. Coupling crop and bio-economic farm modelling to evaluate the revised fertilization regulations in Germany. Agricultural Systems 177, 102687. https://doi.org/10.1016/j.agsy.2019.102687

Kuhn, T., Schäfer, D., 2018. A farm typology for North Rhine-Westphalia to assess agrienvironmental policies (Discussion Paper No. 1), Discussion Paper Series. Institute for Food and Resource Economics, Bonn.

Kuhn, T., Schäfer, D., Holm-Müller, K., Britz, W., 2019. On-farm compliance costs with the EU-Nitrates Directive: A modelling approach for specialized livestock production in northwest Germany. Agricultural Systems 173, 233–243. https://doi.org/10.1016/j.agsy.2019.02.017

Lengers, B., Britz, W., Holm-Müller, K., 2014. What Drives Marginal Abatement Costs of Greenhouse Gases on Dairy Farms? A Meta-modelling Approach. Journal of Agricultural Economics 65, 579–599. https://doi.org/10.1111/1477-9552.12057

Lowenberg-DeBoer, J., Huang, I.Y., Grigoriadis, V., Blackmore, S., 2020. Economics of robots and automation in field crop production. Precision Agriculture 21. https://doi.org/10.1007/s11119-019-09667-5

Martini, D., Herzig, D., Ladwig, G., Kunisch, M., 2014. Semantic search: finding KTBL's planning data and reusing them in IT systems. Landtechnik 69, 12–18.

Reidsma, P., Janssen, S., Jansen, J., van Ittersum, M.K., 2018. On the development and use of farm models for policy impact assessment in the European Union – A review. Agricultural Systems 159, 111–125. https://doi.org/10.1016/j.agsy.2017.10.012

Rose, D.C., Sutherland, W.J., Parker, C., Lobley, M., Winter, M., Morris, C., Twining, S., Ffoulkes, C., Amano, T., Dicks, L.V., 2016. Decision support tools for agriculture: Towards effective design and delivery. Agricultural Systems 149, 165–174. https://doi.org/10.1016/j.agsy.2016.09.009

Schäfer, D., Britz, W., Kuhn, T., 2017. Flexible Load of Existing Biogas Plants: A Viable Option to Reduce Environmental Externalities and to Provide Demand driven Electricity? German Journal of Agricultural Economics 66.

Sørensen, C.G., Fountas, S., Nash, E., Pesonen, L., Bochtis, D., Pedersen, S.M., Basso, B., Blackmore, S.B., 2010. Conceptual model of a future farm management information system. Computers and Electronics in Agriculture 72, 37–47. https://doi.org/10.1016/j.compag.2010.02.003

Sparrow, R., Howard, M., 2021. Robots in agriculture: prospects, impacts, ethics, and policy. Precision Agric 22, 818–833. https://doi.org/10.1007/s11119-020-09757-9

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B.L., Lassaletta, L., de Vries, W., Vermeulen, S.J., Herrero, M., Carlson, K.M., Jonell, M., Troell, M., DeClerck, F., Gordon, L.J., Zurayk, R., Scarborough, P., Rayner, M., Loken, B., Fanzo, J., Godfray, H.C.J., Tilman, D., Rockström, J., Willett, W., 2018. Options for keeping the food system within environmental limits. Nature 562, 519–525. https://doi.org/10.1038/s41586-018-0594-0

Troost, C., Walter, T., Berger, T., 2015. Climate, energy and environmental policies in agriculture: Simulating likely farmer responses in Southwest Germany. Land Use Policy 46, 50–64. https://doi.org/10.1016/j.landusepol.2015.01.028

Vishwanath, R., Moreno, J.F., 2018. Review: Semen sexing – current state of the art with emphasis on bovine species. Animal 12, s85–s96. https://doi.org/10.1017/S1751731118000496

Chapter 2
Economic opportunities of using crossbreeding and sexing in Holstein dairy herds[1]

Abstract

With the increasing availability of sexed semen, farms have the opportunity to select genetically superior dams for producing their replacement animals and produce crossbred calves for beef production of higher economic value from the remainder of the herd. However, higher costs and reduced fertility of sexed semen complicate the decision of when and to which extent sexed semen should be applied in a herd. The objective of this study was to explore the economically optimal sexed semen and crossbreeding utilization among North Rhine-Westphalian dairy farms in a holistic single farm model. For the analysis, a representative sample of farms was derived from Latin Hypercube sampling, based on the observed distribution of farm characteristics from the official North Rhine-Westphalian Farm Structure Survey (FSS) data. Market- and technology-related input parameters such as output prices and sexed semen accuracy and fertility were included in the sampling procedure. Modeling results of the systematic sensitivity analysis were evaluated in a statistical meta-model. We found that the profit maximizing sexed semen and crossbreeding utilization was highly heterogeneous among the farms. Farms with lower stocking densities < 2

---

[1] This chapter is published in the *Journal of Dairy Science* as:

Pahmeyer, C., Britz, W., 2020. Economic opportunities of using crossbreeding and sexing in Holstein dairy herds. Journal of Dairy Science 103, 8218–8230. https://doi.org/10.3168/jds.2019-17354

LU/ha were generally found to produce excess heifers for sale, while farms with stocking densities > 2 LU/ha were producing crossbred calves and using sexed semen only to produce replacement animals. On average, female sexed dairy semen was used on 25.3% of all inseminations. Beef semen (both sexed and conventional) for producing crossbred calves was used on an average of 21.5% of the inseminations. The combination of sexed semen and crossbreeding increased profits from €0 to €568 per cow per year, with an average of €79.42 per cow per year. Farms characterized by low stocking densities (< 2 LU/ha) and above average replacement rates (> 40%) were found to have higher profit increases as a result of selling more heifers from the use of sexed semen. Overall sexed semen and crossbreeding adoption were most sensitive to stocking density and average cow longevity, as well as additional costs for sexed semen and sexed semen accuracy. Our results show the potential of modern breeding technologies to improve dairy farm profits and the need to judge their profitability in the light of farm-specific production settings.

Keywords

sexed semen, crossbreeding, farm model, economics

2.1 Introduction

The average lifespan of cows has slightly increased over the past decades. However, cow vitality and fertility remain important issues in German high-input/high-output dairy herds (Martens, 2016; Römer et al., 2018). On average, Holstein cows are culled after only 2.7 lactations in Germany, with almost 30% leaving during the first lactation (Arbeitsgemeinschaft Deutscher Rinderzüchter, 2018; Römer et al., 2009). As a consequence, herd replacement rates remain high, and without the use of sexed semen almost all female offspring is required for replacements in the milking herd (De Vries et al., 2008).

The potential for economic issues in such a limited lifespan of dairy cows is high. A high replacement rate causes considerable rearing and replacement costs per cow. As a by-product of the necessary purebred female offspring, a high share of purebred male calves per average cow is produced with low weight gains and reduced meat quality compared to their beef breeds counterparts. The male calves are subsequently of relatively low economic value and contribute little to overall farm profits (Wolfová et al., 2007). A high number of rearing heifers relative to the productive herd also implies higher nutrient excretion per productive cow with related costs required in order to comply with environmental legislation, such as the recently revised Fertilization Ordinance (FO) in Germany. There is also higher feed consumption per productive cow, again with detrimental impacts on environmental performance (Weiske et al., 2006).

In the recent past, the application of sexed semen has been growing rapidly in the dairy industry (Holden and Butler, 2018). Sexed semen allows pre-determining the sex of a calf with an accuracy around 90% (Seidel, 2014), which opens new possibilities to address the aforementioned issues. However, in comparison to conventional semen, prices for sexed semen remain higher (Schneichel, 2017), and lower conception rates of ca. 70-90% (DeJarnette et al., 2011; Healy et al., 2013; Maicas et al., 2020) increase the number of required inseminations (Seidel, 2014).

In addition to using sexed semen, farmers may use beef semen in order to produce dairy-beef crossbred calves. The resulting crossbred calves perform better in fattening systems, yielding higher sales value (Wolfová et al., 2007) and improve the overall efficiency of the dairy-beef chain. Combining the use of sexed semen and crossbreeding constitutes a promising alternative to conventional breeding methods. It may improve economic results, deliver a higher proportion of female calves, and allow for selecting only the genetically highest ranking cows for replacements. Crossbred calves that perform better in fattening may be produced from the remaining dams of the herd.

Different aspects of sexed semen with and without crossbreeding have been addressed in the literature. McCullock et al. (2013) studied the influence of key variables such as market, management, and technology on the profitability of using sexed semen in a high yielding Holstein herd in Colorado. Their results suggested that management variables (e.g. conception rates) and the price of dairy heifer calves had a significant effect on the net present value gain per cow, whilst the cost of sexed semen and the milk price showed relatively little effect on profitability. Potential effects on the rate of expansion for heifers and lactating cows in a pasture-based system using sexed semen were modeled by Murphy et al. (2016). Five different breeding strategies were analyzed, showing that the scenario where sexed semen was used on heifers and a targeted group of cows facilitated the fastest possible expansion. A stochastic, bioeconomic spreadsheet model was employed by Cottle et al. (2018) in order to analyze the profitability of using sexed semen in a high-input/high-output dairy herd. Their findings suggested that inseminating both heifers and cows with sexed semen was the most profitable in the simulation. The study emphasizes the relatively elevated effect of pregnancy rate and the genetic value of dairy bulls for determining the financial advantages of sexed semen usage. Ettema et al. (2017) combined the two simulation models SimHerd (Østergaard et al., 2005) and ADAM (Pedersen et al., 2009) to study the hypothesis, that sexed semen increased the genetic gain and overall net return depending on herd management. The hypothesis that the potential for beef semen to increase genetic level would be herd-specific was supported by their findings. However, they concluded that none of the scenarios modeled were profitable under Danish circumstances when the value of the increased genetic level was not included.

Given that the existing literature found large differences regarding the driving factors for the profitability of sexed semen and beef semen use for different types of dairy farms, a proper assessment should incorporate a population of farms and their characteristics for the evaluation of profitability. However, a detailed whole-farm analysis studying the profit maximizing shares of sexed dairy semen, sexed beef semen and conventional beef semen (crossbreeding) depending on farm-characteristics is still missing. The objectives of this paper were therefore three-fold. First, we determined the extent to which dairy farms in our study region would use either sexed dairy semen, sexed beef semen, conventional beef semen or a combination of these under profit-maximizing behavior. Second, we explored factors explaining differences in the adoption rates of sexed dairy semen and beef semen (both sexed and conventional beef semen) between different farms or market conditions or both. Third, we studied the potential economic benefits for different types of dairy farms and the extent to which they rely on market conditions.

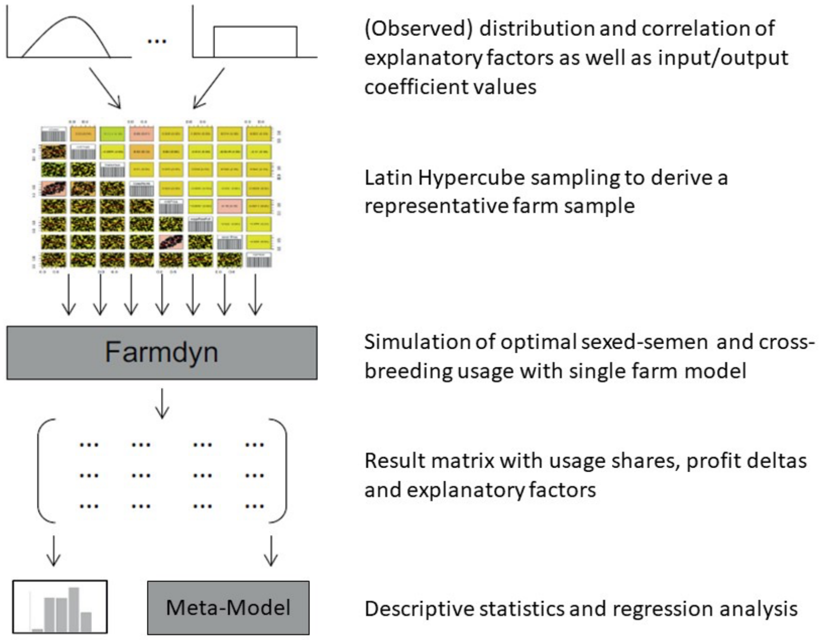


Figure 2.1 Overview of modeling framework, adapted from Kuhn et al. (2019) and Lengers et al. (2014).

2.2 Materials and Methods

A three-step modeling framework was used to assess the profit maximizing shares of sexed semen and beef semen use (Figure 2.1).

The method was largely based on a meta-modeling approach proposed by Lengers et al. (2014) and Kuhn et al. (2019).

At first, a representative sample for the farm population including draws for input and output prices and sexed semen characteristics was generated by Latin Hypercube Sampling (LHS). The sampling was based on data from official agricultural statistics. In a second step, each of the sample farms was simulated in the single farm optimization model FarmDyn. In order to depict the economic effects of using sexed dairy semen and (sexed) beef semen, the model was solved two times for each farm. First, the model was solved without the possibility to use sexed dairy semen and beef semen for crossbreeding (baseline). In a second run, both sexed semen and beef semen (both conventional and sexed beef semen) were made available. Note that the share of sexed dairy, sexed beef, and conventional beef semen was an endogenous variable in our optimization model such that the model would simulate optimal usage intensity (which may also be zero) to address the research questions. In a third step, a statistical meta-model was derived to explain overall sexed semen/beef semen share and profit deltas resulting from sexed semen and beef semen usage.

2.2.1 *Sampling procedure*

Our analysis examined specialized dairy farms in the German federal state of North Rhine-Westphalia (NRW). Roughly 10% of the German dairy population is based in NRW, with more than 4,300 farms specialized on dairy production (Statistisches Bundesamt, 2018). Selling their male calves two weeks after birth, these farms generally produce their own replacement heifers. Additional heifers might be sold for replacements or slaughter. Distributions reflecting single farm data for the study region were used from the German Farm Structure Survey (FSS) 2016 as reported by Kuhn et al. (2019). The data covered factors regarding farm endowments, such as farm sizes, grassland shares, stocking densities, and manure storage capacities as explanatory factors. Regional data regarding crop production, such as maize silage and grassland yields, were taken from KTBL (2019). Ranges of input coefficients for animal prices, milk yields, lactation lengths, and first calving ages were derived from KTBL (2018). Data on sexed semen accuracy, and additional insemination efforts (IE) when using sexed semen (sexed semen conception rate) were drawn from Seidel (2014) and Butler et al. (2014). The explanatory factors and their ranges are depicted in Table 2.1. As empirical distributions for input parameters not covered in the FSS were not available, uniform distributions for these parameters were assumed. The correlation between these parameters was assumed to be zero, reducing the risk of multi-collinearity in the statistical meta-model. In order to depict the whole value range of the model input parameters, a sample of the farm population of ca. 1700 farms was generated using LHS (McKay et al., 1979).

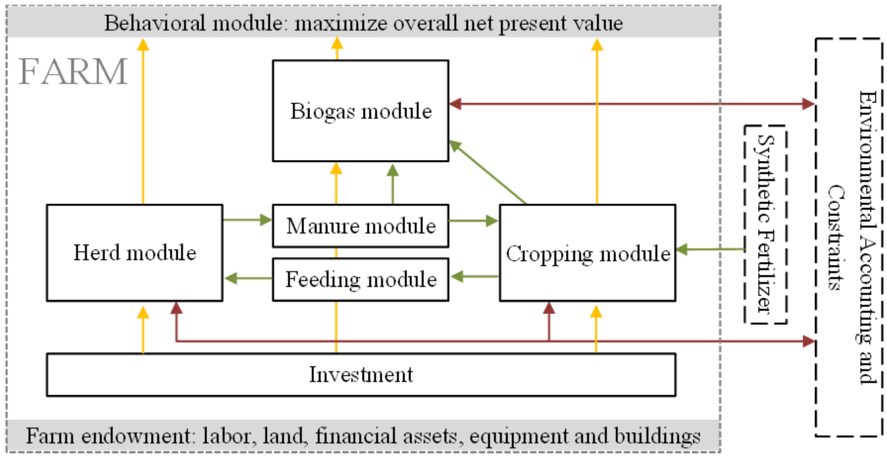
Explanatory factor	Minimum	Median	Maximum	Data source
Farm size [ha]	8.14	61.24	221.35	FSS 2016 in Kuhn et al. (2019)
Grassland share [%]	6	51	100	FSS 2016 in Kuhn et al. (2019)
Grassland yield [t DM/ha]	6.19	6.94	7.69	KTBL (2019)
Maize Silage yield [t FM/ha]	42.9	46.55	50.2	KTBL (2019)
Stocking density [LU/ha]	0.63	1.75	5.94	FSS 2016 in Kuhn et al. (2019)
Milk Yield [kg ECM/a]	6800	9400	12000	KTBL (2018)
Avg. cow longevity [lactations/cow]	1.7	4.15	6.6	KTBL (2018)
Calving interval [d]	365*	408.5	452	KTBL (2018)
Additional costs female sexed semen doses [€/n]	13	24.5	36	Besamungsverein Neustadt (2019)
Additional insemination effort sexed semen [n]	0	0.575	1.15	Own calculation based on Butler et al. (2014)
Sexed semen accuracy [%]	75	87.5	100	Seidel (2014)
Price heifer [€/head]	1040	1490	1940	KTBL (2018)

Beef bull calf [€/head]	70	195	320	KTBL (2018)
Beef heifer calf [€/head]	40	145	250	KTBL (2018)
Dairy heifer calf [€/head]	10	50	90	KTBL (2018)
Dairy bull calf [€/head]	10	80	150	KTBL (2018)

Table 2.1 Characterization and sources of explanatory factors.*The minimum calving interval was set to 365 days in order to fit the assumptions made in the materials and methods section

2.2.2 *Farm modeling*

For the farms derived from the sampling procedure, the adoption and economic gain of sexed semen and crossbreeding were estimated. For the analysis, we enhanced the dynamic, mixed-integer linear programming model FarmDyn (T. Kuhn et al., 2019a; Lengers et al., 2013, 2014) to incorporate adequate representations of herd dynamics and breeding techniques. FarmDyn is an open-source and open-access bio-economic single farm model written in the General Algebraic Modeling Language (GAMS Development Corporation, 2019). For further technical information regarding the model FarmDyn, see the technical documentation (Britz et al., 2016). Its modular structure (illustrated in Figure 2.2) allows simulating different farming types, including dairy, suckler, beef, and arable farms. Indivisibilities in investments and labor use, such as buildings and machinery, are captured by the use of integer variables (mixed-integer linear programming (MILP)). Farm management decisions, e.g. feeding, use of sexed semen or crossbreeding, manure management and labor distribution are modeled with a monthly resolution (partially bi-weekly). The model is parameterized for German conditions with the use of detailed farm planning data provided by KTBL (2018) and LfL Bayern (2018). Model results were validated with individual project case study data.



Remark: — Represents mass transfers from one module to another
— Represents monetary transfers
— Represents environmental and related transfers.

Figure 2.2 Simplified flowchart illustrating the main modules of the model FarmDyn. Source: Britz et al. (2016).

For the current study, the deterministic comparative static version of the model was utilized which maximizes the net present value (NPV) over a predefined planning horizon of farms using farm-household resources (labor, land, financial assets). A rational, fully informed, and risk-neutral decision maker was assumed. We refrained from modeling risk and risk behavior as we lacked any defendable source to parameterize risk behavior in the required detail. Moreover, as no closed-form determination of variance-covariance matrices for the decision variables was possible, the application of a mean-variance analysis (Markowitz, 1952) was not feasible. The influence of variability in input coefficients across farms on the overall NPV and other factors was analyzed by a structured sensitivity analysis.

As investment decisions were not at the core of our analysis, we performed a comparative-static analysis where continuous reinvestments for machinery and sunk costs for housing were assumed and related costs annualized. Investments in new housing were disabled in order to capture the short-term effects of sexed semen and crossbreeding usage. Consequently, herd dynamics were depicted in a steady-state model as described in the following section such that impacts of the advanced breeding methods on an average production year were depicted. The model was constrained by available resources, possible production processes, allowed crop rotations, off-farm working opportunities and two restrictions relating to agri-environmental legislation. These were the German FO as the implementation of the European Nitrates and Water Framework Directives and the greening obligations under the First Pillar of the Common Agricultural Policy.

The dairy module characterized one of the possible farm branches in the model. Here, mass flows as feeds, manure, and animals were described, and linked to the economic optimization part of the model. Feed rations were optimized endogenously according to the nutrient and dry matter constraints defined by the Zifo2 feed optimization application (LfL Bayern, 2016). Also, different feeding regimes such as grazing, partial grazing or non-grazing were accounted for, depending on the endowments present on the farm. Manure handling was assumed to be outsourced to a contractor and may have been spread on own fields or, with additional costs, exported to other farms. Land leasing and buying of forages were disallowed in order to depict the short-term effects of sexed semen and crossbreeding usage only.

The model differentiated between calf-, heifer-, and cowherds of different sex and ages. Each animal of a herd represented an average animal of the herd in the model. Therefore, production parameters such as milk yield, lactation profile, breeding value, number of services, and pregnancy rate were equal among animals within each herd group.

Default modeling parameters were specified in order to reflect the range of dairy farms in the study area. Milk yields between 6,800 up 12,000 kg ECM were realized (KTBL, 2018). The roughage intake was assumed to be at least 60% of the total dry matter intake of the cows. Cows were not given access to pastures, and year-around calving was assumed. The availability of grass and maize silage per cow as an endogenous variable reflected the individual farms' endowments (as mentioned above).

For heifers, an average number of 1.6 services, and for cows 2.3 services per pregnancy were assumed when conventional semen was used, implying a fixed heat detection rate of 70% (Römer et al., 2013). Cows and heifers being culled due to reproductive (exceeded maximum number of services) or nonreproductive (e.g. lameness, mastitis) reasons were implicitly captured by an exogenously assumed average cow longevity in the herd. In addition, an annual mortality rate of 5% was applied to the model (KTBL, 2018).

Depending on the specific farm characteristics, sufficient replacement heifers were reared targeting 23 to 25 months as age of first calving. Depending on the age of first calving, feed composition, labor requirements and final weight were adapted accordingly. Surplus calves were assumed to be sold to the market at < 1 month of age.

2.2.3 *Implementation of sexed semen and crossbreeding*

In addition to conventional dairy semen, the opportunity to use sexed dairy, sexed beef, or conventional beef semen was introduced into the above-described farm model as an endogenous decision.

For the conventional and sexed beef semen, it was assumed that Belgian Blue sires were used for crossbreeding. Costs for conventional beef semen were assumed to be equal to conventional dairy semen in the model. Similarly, sexed beef semen was assumed to be equal in costs and fertility as sexed semen of a Holstein sire. Crossbred calves were assumed to be sold right after birth, additional on-farm fattening was not considered.

Regarding sexed semen (both dairy and beef sexed semen), it was assumed that when sexed semen was used on the first service, it was used throughout the remaining services until a pregnancy was achieved. The conception rate of sexed semen for both heifers and cows varies in the analysis (as mentioned above) and was assumed to be 75 – 100% of the conception rate of unsorted conventional semen (Butler et al., 2014). In order to determine the effect of the reduced conception rate of sexed semen on the average number of services required for a pregnancy (IE), conception rates depending on the service number by Kuhn et al. (2006) were evaluated for heifers, while for cows a fixed conception rate of 35% was assumed for conventional unsorted semen. It was assumed that for a successful insemination using sexed semen, the same cumulative probability that any of the inseminations have been successful as with conventional unsorted semen would be required. Given the average of 1.6 services for heifers, and 2.3 services for cows, an additional average increase of 0 – 1.15 services (again varied in the analysis) was derived when sexed semen was used. This way, the higher conception rates of using sexed semen on heifers compared to multiparous cows were reflected in the model. Herd specific parameters that vary in the analysis are displayed in Table 2.1, while default static input values are presented in a supplementary material file (Supplemental Table S1; https://doi.org/10.3168/jds.2019-17354).

The additional number of services when using sexed semen had a number of implications on the model: On heifers, the average first calving age was (involuntarily) shifted upwards when sexed semen was used. Consequently, additional feed and labor were required, leading to an increase in rearing costs. On multiparous cows, the involuntarily extended lactation length led to a change in the lactation profile and fewer calvings per year. For both heifers and cows, the additional input costs per successful sexed semen insemination were assumed to be the total IE multiplied by the price premium of sexed semen. Further technical aspects of the specific calculation of herd sizes and calving distributions within the modeling framework are presented in the Supplementary Material; https://doi.org/10.3168/jds.2019-17354.

Among other impacts, the use of sexed semen and beef semen has an effect on dystocia (Norman et al., 2010). In order to depict the effect of differing dystocia risk among heifers and cows regarding bull calves, heifer calves, and beef calves, we used dystocia risk scores and related costs as described by McCullock et al. (2013) (Supplemental Table S1; https://doi.org/10.3168/jds.2019-17354).

Literature has shown the importance of genetic improvement as an economic factor for the use of sexed semen or beef semen or both (Cottle et al., 2018; Ettema et al., 2017; McCullock et al., 2013). In general, genetic improvement is obtained by mating animals so that their offspring is genetically superior compared to the population in terms of the breeding goal (De Vries, 2017). Through the use of sexed semen, genetically superior animals can be selected to produce the next generation of replacement heifers, thus improving the genetic level of the herd. To increase the genetic level even further, beef semen can be used on genetically inferior multiparous cows, so that their offspring does not enter the milking herd. We based the calculation of genetic merit on the results outlined by Ettema et al. (2017). Similar to their approach, we valued one genetic standard deviation of the Nordic Holstein breeding goal with €89 per cow-year. Improvements in genetic level for using sexed dairy semen on heifers and beef semen (sex sorted and unsorted) cows were derived from their results by linear regression. Breeding 1% of all heifers with sexed semen resulted in an estimated genetic increase of ca. 0.001 standard deviation units of the breeding goal while breeding 1% of all multiparous cows with beef semen resulted in a genetic increase of ca. 0.0017 standard deviation units of the breeding goal respectively. The resulting economic value of differences in the genetic level induced by sexed semen and beef semen use (genetic return) was then calculated by multiplying the difference in genetic standard deviation units by €89.

2.2.4 *Statistical meta-modeling*

In order to quantify which factors significantly impact the adoption of sexed semen, beef semen, as well as the profit gain resulting from sexed semen and beef semen use, a statistical meta-model was constructed for each of the three cases. The meta-models approximated the input and output transformations of the FarmDyn model, resulting in a simplified statistical model summarizing all simulations runs of the sample population. The meta-models were specified as multiple linear regression models, where the explanatory factors displayed in Table 2.1 were defined as the independent variables, and the share of sexed semen, beef semen, and profit gain from sexed semen and beef semen use as the dependent variables respectively. The statistical analysis was conducted with the software R (R Core Team, 2017).

2.3 **Results**

We found the profit maximal sexed semen and crossbreeding application among the studied farm population to be highly heterogeneous, reflecting the distribution of farm characteristics and input coefficients described in Table 1. Under these assumptions, 93.3% of the farms were simulated to use female-sexed dairy semen (X-chromosome enriched) on heifers at least once. Furthermore, 6.4% of the farms were simulated to use male-sexed beef semen (Y-chromosome enriched) on heifers at least one time. Female-sexed dairy semen used on cows was found to be profit maximal in 6.9% of the farms to some extent. Male-sexed beef semen on cows was found to be utilized in 0.1% of all farms. Beef semen use, both sexed and conventional, was found to be used in 49% of the farms for at least one insemination. On the farms that used beef semen, an average of 76.2% of the cows and 13.3% of the heifers were bred with beef semen.

Female-sexed dairy semen was used on an average of 25% of all inseminations in the optimum. An average of 66% of all heifers were bred with female-sexed dairy semen, and 5% of all cows. The average herd replacement rate was 22%, with an average cow longevity of 4.46 lactations per cow among the sample population.

Using beef semen for crossbreeding (both conventional and sexed semen) was found to be profit maximal for an average of 21% of all inseminations. Conventional beef semen was used on an average of 37% of all cow inseminations, and 6.5% of all heifer inseminations.

On average 2% of all heifers and less than 1% of all cows were bred with sexed beef semen in the simulation results.

2.3.1 *Drivers of overall sexed semen and crossbreeding usage*

In order to identify the main drivers of these profit maximal shares of sexed semen and crossbreeding usage, the statistical meta-model described in the previous section was analyzed. The standardized regression coefficients for the explanatory variables are displayed in Table 2.2. While many of the explanatory variables were found to have a significant impact on the dependent variables, a few explain a greater part of the variance. This was expected given the high number of observations and the absence of measurement and reporting errors. In the case of the relative share of female-sexed dairy semen used on all inseminations on a farm, the β-coefficients of management factors such as cow longevity, as well as technology factors such as conception rates (expressed through additional required IE), and general farm endowment factors as the overall stocking density were found to have the largest absolute impact on the overall profit maximal adaptation.

	Dependent variable (β-coefficients):		
	Share sexed dairy semen [all inseminations]	Share beef semen [all inseminations]	Profit delta sexed semen and crossbreeding
	(1)	(2)	(3)
Cow longevity [n lactations/cow]	-0.471^{***}	0.091^{***}	-0.307^{***}
Additional IE sexed semen [n]	-0.418^{***}	-0.070^{***}	-0.287^{***}
Stocking density [LU/ha]	-0.307^{***}	0.689^{***}	-0.455^{***}
Calving interval [d]	-0.279^{***}	-0.047^{***}	-0.188^{***}
Dairy bull calf [EUR/head]	-0.073^{***}	-0.079^{***}	-0.070^{***}
Beef bull calf [EUR/head]	0.062^{***}	0.124^{***}	0.106^{***}
Additional costs female sexed semen doses [EUR/n]	-0.047^{***}	-0.010	-0.037^{**}
Milk Yield [kg ECM/a]	-0.046^{***}	0.093^{***}	-0.032^{*}
Price heifer [EUR/head]	0.032^{**}	-0.048^{***}	0.220^{***}
Grassland share [%]	0.025^{*}	-0.066^{***}	0.081^{***}
Sexed semen accuracy [%]	0.022	0.027^{*}	0.076^{***}

Farm size [ha]	0.014	-0.079***	-0.133***
Beef heifer calf [EUR/head]	0.013	0.167***	0.080***
Maize Silage yield [t FM/ha]	0.012	0.002	0.026
Manure storage capacity [m]	-0.010	-0.013	-0.010
Dairy heifer calf [EUR/head]	-0.005	-0.001	-0.001
Grassland yield [t DM/ha]	-0.003	-0.014	0.019
Constant	-0.000	-0.000	-0.000
Observations	1,736	1,736	1,736
R^2	0.616	0.621	0.505
Adjusted R^2	0.613	0.617	0.500
Residual Std. Error (df = 1718)	0.622	0.619	0.707
F Statistic (df = 17; 1718)	162.442***	165.411***	103.150***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2.2 Multiple linear regression model output (meta-model of the single farm model FarmDyn) on overall sexed semen share, beef semen share and profit deltas resulting from sexed semen and crossbreeding usage (dependent variables) for North Rhine-Westphalian dairy farms

The top half of Figure 2.3 shows the simulated relative share of female sexed dairy semen used on the farms, depending on the farms' individual stocking density and average cow longevity. In addition, Supplemental Figure S2; https://doi.org/10.3168/jds.2019-17354 displays the same graph, however showing the sexed semen conception rates (expressed as additional insemination effort) on the color axis. From the two graphs, it can be seen that very high sexed semen utilization shares up to 100% of all insemination were found to be profit maximal for farms with stocking densities < 1.5 LU/ha, combined with below average cow longevity (< 3 lactations per cow), and high sexed semen conception rates (> 90% of conventional semen conception rate, requiring fewer additional inseminations when sexed semen was used). As indicated by the β-coefficient displayed in Table 2.2, herds with high average cow longevity (> 3 lactations per cow) were in general found to be utilizing sexed semen significantly less than farms with below average cow longevity. Farms using sexed semen on few inseminations were mostly characterized by higher stocking densities (> 2 LU/ha), combined with above average cow longevity (> 4 lactations per cow), and low sexed semen conception rates (< 80% of conventional semen conception rate).

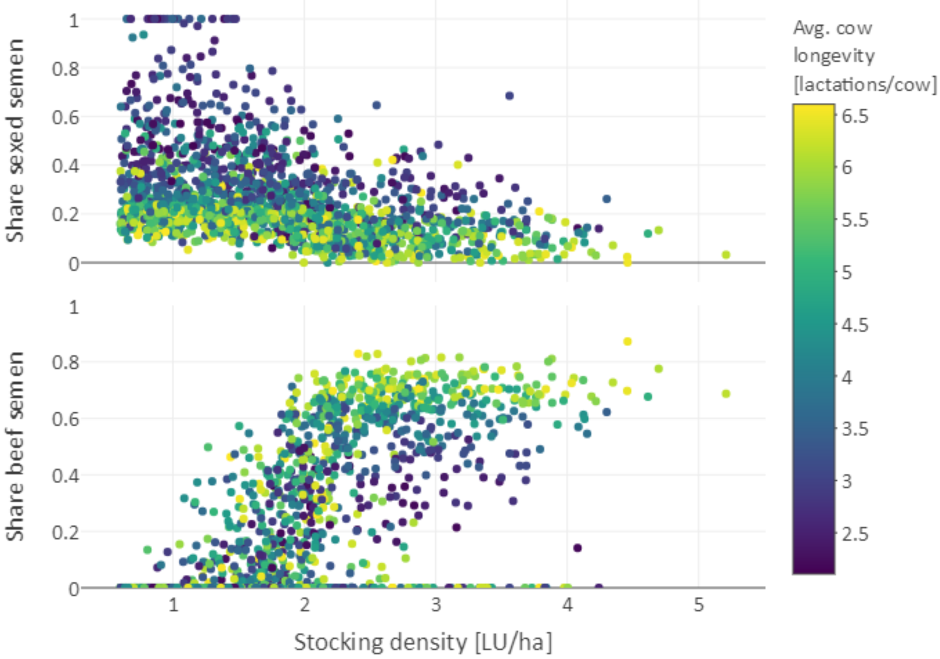


Figure 2.3 Profit maximizing share of female sexed dairy semen used on all inseminations (upper part), compared to the share of beef semen (sum of sexed beef semen and conventional semen) used on all inseminations (bottom part) in the North Rhine-Westphalian study population. Each dot represents a farm in the sample population. LU = livestock units.

The bottom half of Figure 2.3 displays the simulated relative share of beef semen (sexed and conventional beef semen jointly considered) used on the farms for producing crossbred calves, again depending on the farms' individual stocking density and average cow longevity. Farms with stocking densities > 2 LU/ha, together with average cow longevities > 5 lactations per cow were simulated to have high uptakes of beef semen mostly ranging from 70-80% of all inseminations in the optimum. Farms with lower stocking densities < 2 LU/ha or lower than average cow longevity (< 3 lactations per cow) or both, on the other hand, were simulated to have a wider range of beef semen uptake between 0-60%. As displayed by the β-coefficients in the “Share beef semen” column of Table 2.2, apart from the stocking density a multitude of market factors such as beef heifer calf and beef bull calf prices had a positive impact on overall beef semen uptake, while overall farm size (ha), dairy bull and heifer prices, as well as sexed semen conception rates, had a negative impact on the overall uptake. When beef semen was used in a farm, the share of sexed beef semen on all beef semen used was ranging from 0-100% among the simulated population. On average, 2.7% of the beef semen used was sexed beef semen.

2.3.2 *Effects on farm profitability*

The potential economic benefits of using sexed semen and crossbreeding were again found to be highly heterogeneous within the North Rhine-Westphalian dairy farm population, ranging from €0 to €568/cow and year. On average, farms in the population sample could increase their profits by €79.42/cow and year (median €59.52 per cow and year) by applying sexed semen, crossbreeding or both. Multiplied with a simulated average herd size of 91 cows, the mean sexed semen and crossbreeding induced profit increase would add up to a total of ca. €7,225 per farm per year. When only sexed and conventional dairy semen (no crossbreeding) was made available to the farms, the average profit increase was €65.35 per cow (median €40.44 per cow) compared to the baseline. In the run where only conventional beef semen and conventional dairy semen (no sexed semen) was made available to the farms, the average profit increase was €10.06 per cow (median €8.33 per cow), with a range of €0 to €227 per cow among the sample population, reflecting the possibility to use crossbreeding.

Table 2.2 reports on a meta-model that analyzed these profit differences. The variable coefficients in the third column indicate that the economic benefits of utilizing sexed semen and crossbreeding were driven by market factors such as heifer, beef bull calf and beef heifer calf prices. On the other hand, increasing stocking densities, as well as cow longevity and lower sexed semen conception rates were found to diminish additional profit gains compared to the baseline, where neither sexed semen nor crossbreeding was used on the farms.

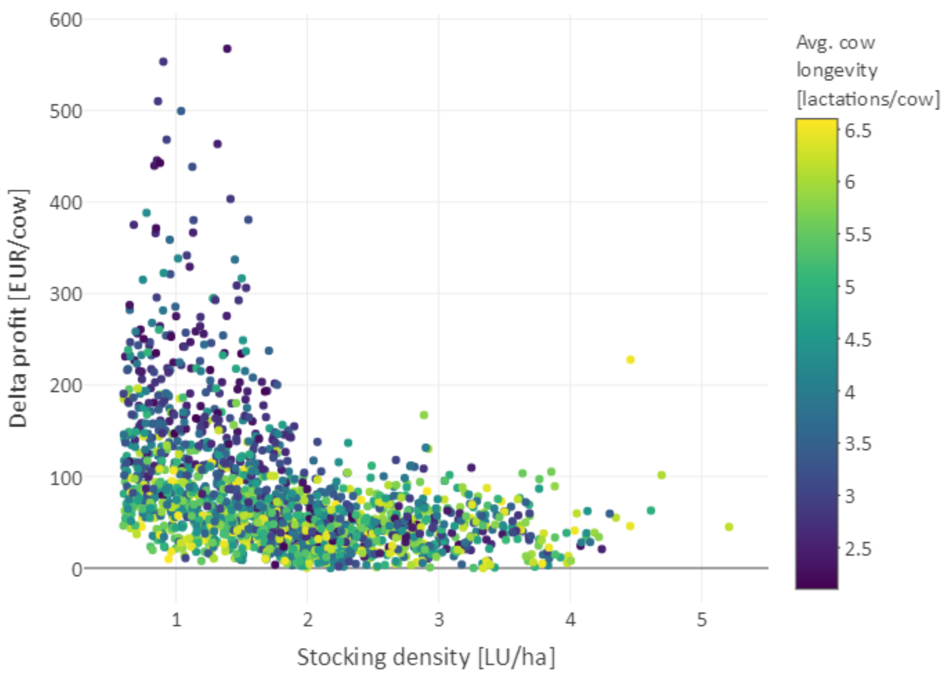


Figure 2.4 Simulated profit increase (€ per cow) induced by the profit maximal use of sexed dairy semen and sexed/conventional beef semen among the North Rhine-Westphalian dairy farm population. Each dot represents a farm in the sample population.

Figure 2.4 shows additional profits per cow when sexed semen and sexed/conventional beef semen were made available to the farm population, compared to the baseline. The largest profit increases of more than €500 per cow were observed in farms where below average stocking densities (< 2 LU/ha), together with below average cow longevity (< 3 lactations per cow) were prevalent. On the other hand, farms with stocking densities > 2 LU/ha and high average cow longevity (> 5 lactations per cow) were most often found to have profit gains in the range of €0 to €30 per cow.

The economic impact of the change in genetic level induced by sexed semen and beef semen use (genetic return) was ranging from €0 to €21 per cow, with an average genetic return of €4.98 per cow (median €2.16 per cow) among the sample population. Changes in costs linked to dystocia were ranging from €3.80 per cow to €12.45 per cow, with an average change of €0.31 per cow.

2.4 **Discussion**

2.4.1 *Results*

Our results imply that a profit-maximizing utilization of crossbreeding and sexed semen among North Rhine-Westphalian dairy farms could improve profits from €0 to €568 per cow per year. We found that almost half of the simulated farm population could increase their profits by using female-sexed dairy semen on heifers for replacement heifer production and using beef semen on cows to some extent. However, potential economic gain of sexed semen and crossbreeding utilization, as indicated in the results section, varied greatly depending on individual farm endowments such as stocking density and average cow longevity, among other factors. In line with the findings of McCullock et al. (2013), the strategy of producing excess heifers for sale with the use of sexed dairy semen was found to be profit-maximizing, though only for farms with below average stocking densities (< 1.5 LU/ha). Especially when a high grassland share was present, rearing excess heifers was a profitable opportunity for these farms, as often no production alternative for grassland was present. The overall highest potential economic gains of up to €568 per cow were observed when below average stocking density, an average cow longevity < 3 lactations per cow, high heifer prices and favorable sexed semen parameters (high conception rates and accuracy of sexed semen) were jointly present. For these farms, using sexed semen on all heifers, as well as all cows was found to be most profitable, confirming the findings of Cottle et al. (2018) for such types of farms.

Farms with higher average cow longevities were generally rearing excess heifers even before the introduction of sexed semen into the model, and thus had a smaller profit gain from the new technology. Given increased profits per cow, many of these farms would have potentially invested in new cattle housing in order to further expand their herd in the model. However, as stated in the previous sections, investments in new cattle housing were disabled in order to depict the rather short-term effects of sexed semen and crossbreeding usage.

With increasing stocking densities, feed competition among animal groups within the farm increased. As buying of roughages from the market was disallowed in the analysis, farms that could barely sustain their cow herd size due to limited own fodder production had little incentives to produce excess heifers. In order to rear excess heifers, these farms would have needed to reduce their cow herd, increasing the marginal production costs of heifers by the opportunity costs of the cows that could have been fed instead. Farms with stocking densities > 2 LU/ha and milk yields > 10,000 kg ECM were therefore rarely found to produce more heifers than required for their own replacements. These findings support the results of Ettema et al. (2017), where the profitability of excess heifers was largely determined by additional heifer rearing costs.

Farms with higher stocking densities were often found to be using beef semen in order to produce crossbred calves instead of excess heifers. As the crossbred calves were assumed to be sold right after birth, these animals were not competing for additional feed with any additional heifers raised on farm. This way, fodder could be valorized by the cow herd with higher marginal returns.

Farms with average cow longevity well above the mean (> 5 lactations per cow), and limited roughage availability (stocking densities > 2 LU/ha) showed the highest beef semen uptake shares of 60 to 80%. Within this particular group, farms exposed to higher crossbred calf prices (> €200/head) and sexed semen conception rates of > 90% of conventional semen conception rates were particularly likely to show high rates of sexed beef semen usage. Due to the relatively low replacement rates induced by the high average cow longevity, these farms were able to take specific advantage of crossbred calf price premiums. In scenarios with high prices for crossbred calves, farms with limited fodder availability (stocking densities > 2 LU/ha) were generally producing their replacement animals by using sexed semen on the genetically superior heifers in order to be able to produce more crossbred calves. Farms with average cow longevity < 5 lactations per cow together with stocking densities > 2 LU/ha were most often using a similar breeding strategy. However, due to their higher demand for female calves for replacements, the average beef semen uptake of these farms was clearly lower. Among the farms that used beef semen in order to produce crossbred calves, sexed beef semen was found to only play a minor role. On average, only 3.3% of all heifers, and 0.1% of all cows were inseminated using sexed beef semen in the part of the sample that used beef semen.

As the change in genetic level due to sexed semen and beef semen use was simulated based on the results outlined by Ettema et al. (2017), the range of genetic returns from €0 to €21 per cow reflected their findings. On average, the genetic return made up for 13.10% of the sexed semen and beef semen induced profit gain. These results indicate the importance of incorporating the genetic return when assessing the profitability of sexed semen and beef semen use.

Crossbreeding remained profitable for almost half of the farms within the sample population to some extent, even when increased dystocia risk and related costs were considered. However, as farms were able to reduce dystocia risk by using female-sexed dairy semen (especially on heifers), the increased dystocia risk when using beef semen on cows was partially offset.

Despite the presented findings, sexed semen and crossbreeding played a minor role in North Rhine-Westphalian dairy production systems at the time of writing. As noted previously, sexed semen was used in only 6% of all Holstein heifer inseminations in Germany in 2018 (Arbeitsgemeinschaft Deutscher Rinderzüchter, 2018). Figures for multiparous cow inseminations with sexed semen and beef semen inseminations were not available. The method of sexed semen favorable heifers and cows for replacements and the remaining herd with beef semen has been proposed in a series of agricultural magazine articles in Germany (Elite Magazin, 2009; Thomsen, 2016). The discrepancy between its profitability found in research and the limited uptake in practice asks for further research, for instance, considering additional transaction costs for marketing crossbred calves or looking into differences between perceived risks by farmers and the risks assumed in studies.

2.4.2 *Methodological approach*

Our approach extends the existing literature in multiple aspects. Firstly, we applied a highly detailed farm-scale model instead of a model at process scale in order to consider the impacts of varying stocking densities. Secondly, we drew on empirical distributions for the dairy farm population in our study region, and, thirdly, performed systematic sensitivity analysis as well as post-model statistical analysis. The approach could be seen as a more general methodology to estimate potential adaptation rates in a farm population for a technology for which technological parameters and related costs and benefits can be derived from literature, but sufficient farm observations on adaptation are (not yet) available.

Using a holistic single farm optimization model for the analysis has a set of advantages over simulation approaches where fewer variables are endogenous. The FarmDyn model endogenously optimizes multiple decision variables simultaneously. These variables include herd entry and exit dates, fodder production and use of concentrates, grassland management, manure storage, and management, allocation of labor to cash crops and herd management, stable and machinery utilization, as well as inputs required for crop production. Instead of a simplified simulation approach, where levels of (certain) decision variables (e.g. feed uses, heifer breeding strategies, and crop allocation) are pre-determined in scenarios, we simulate profit-maximizing levels. The resulting profit-maximizing strategies for crossbreeding and sexed semen are therefore showing the full potential of the considered options, as they incorporate the complexity of the decision on the whole farm level. While previous studies reported on important drivers of sexed semen (and partially beef semen) adoption for specific farm types given fixed farm endowments (Cottle et al., 2018; Ettema et al., 2017; McCullock et al., 2013), our approach highlights the importance of studying the whole variety of farms in a population.

However, considering more endogenous variables comes at the price of a more complex model, introducing more assumptions and possibly uncertain parameters. It also might mean reducing model detail in other aspects. For instance, in an effort to reduce complexity FarmDyn only models an average animal in each herd, instead of individual animals of a certain genetic level.

As previously stated, empirical distributions for parameters not present in the FSS were not available. For reasons of simplicity, but without loss of generality, uniform distributions without correlations were assumed for these input parameters. Recent data by KTBL (2018) report a spread of the avg. cow longevity between 1.7 and 6.6 lactations per cow (see Table 2.1) which results in an in-going sample mean of 4.15 lactations. This is higher than the average of 2.7 lactations stated by Römer et al. (2009), which could not be incorporated in the LHS approach due to missing information on the underlying distribution.

Upon availability, future research should incorporate observed distributions for all endogenous model parameters in order better reflect the underlying statistical population.

The combination of the LHS with the economic model endogenously removes draws from the LHS which are implausible from an economic viewpoint. Specifically, for 13.2% of the farms in the LHS sample the economic model found that herding cattle was not profitable. In these cases, no cow herd and related profit maximal management choices such as rates of sexed semen use can be observed. Not surprisingly, many of such dropped observations comprise farms with quite high replacement rates. The analyzed sample with positive profits, only, still contains the entire input range of the replacement rates. However, the avg. cow longevity increased from 4.15 to 4.46 lactations in the usable observations, matching a replacement rate of approx. 22%.

Under profit maximization, the model will switch to production alternatives compared to the status quo even if the additional profit gain is marginal. To give an example, the model would choose to produce crossbred calves as soon as their relative profitability surpasses that of purebred ones, as long as other constraints (required replacement heifers, available housing places, etc.) are met. However, in these cases where one production option is found to be only slightly more profitable than another, farmers may decide to remain with their "traditional" production portfolio due to personal preference, additional hidden costs, or other unknown factors. As the average sexed semen and crossbreeding induced relative profit increase was found to be €79.42/cow and year, the effect can be seen as substantial enough to be considered an advantageous production option.

With the relatively high simulated share of crossbred animals produced within the farm population, feedback of the supply increase on the producer price of crossbred calves as discussed by De Vries et al. (2008) could be expected. That feedback cannot be considered in FarmDyn as a supply-side model characterized by exogenous input and output prices. Here, also the systematic sensitivity analysis (meta-model) does not help. Market models such as (partial) equilibrium models incorporate such market feedback by design but miss the detailed depiction of technical production processes required by the present study.

2.5 Conclusion

The profit-maximizing sexed semen and beef semen (for crossbreeding) utilization of North Rhine-Westphalian dairy farms were found to be highly heterogeneous. Farms with lower stocking densities were maximizing profits using sexed semen in order to produce excess heifers for sale, while farms with higher stocking densities were instead producing crossbred calves for sale and using sexed semen on heifers in order to produce replacement animals. Furthermore, sexed semen and crossbreeding usage was found to depend on farm characteristics such as average cow longevity, sexed semen-related parameters such as sexed semen conception rate and accuracy, as well as market factors such as the prices of replacement heifers and crossbred calves. Due to continuous improvements to the sex-sorting technology, and the economic benefits found in our analysis, sexed semen adoption is likely to increase further among the study population. Our results highlight the importance of studying a whole variety of farms in a study population, as driving factors for sexed semen and beef semen adoption were shown to differ substantially among farms.

2.6 Acknowledgements

This research is funded by SusAn, an ERA-Net co-funded under European Union’s Horizon 2020 research and innovation program (www.era-SusAn.eu), Grant Agreement n°696231, and co-funded by the Federal Office for Agriculture and Food (BLE). The authors thank Andreas Beel (Institute for Computer Science, Bonn) for his valuable comments on the binomial expansion of the calving distribution, as well as Till Kuhn and Katie Jarvis for their helpful contribution to proofreading.

2.7 References

Arbeitsgemeinschaft Deutscher Rinderzüchter, 2018. Rinderproduktion in Deutschland 2017 - Ausgabe 2018. Bonn.

Besamungsverein Neustadt, 2019. Preisliste [WWW Document]. URL https://www.bvn-online.de/de/service/rind/besamung/gebuehren-2461.html (accessed 6.17.19).

Britz, W., Lengers, B., Kuhn, T., Schäfer, D., 2016. A highly detailed template model for dynamic optimization of farms. Institute for Food and Resource Economics, University of Bonn. Model Documentation.

Butler, S.T., Hutchinson, I.A., Cromie, A.R., Shalloo, L., 2014. Applications and cost benefits of sexed semen in pasture-based dairy production systems. animal 8, 165–172. https://doi.org/10.1017/S1751731114000664

Cottle, D.J., Wallace, M., Lonergan, P., Fahey, A.G., 2018. Bioeconomics of sexed semen utilization in a high-producing Holstein-Friesian dairy herd. Journal of Dairy Science 101, 4498–4512. https://doi.org/10.3168/JDS.2017-13172

De Vries, A., 2017. Economic trade-offs between genetic improvement and longevity in dairy cattle. Journal of Dairy Science 100, 4184–4192. https://doi.org/10.3168/jds.2016-11847

De Vries, A., Overton, M., Fetrow, J., Leslie, K., Eicker, S., Rogers, G., 2008. Exploring the Impact of Sexed Semen on the Structure of the Dairy Industry. Journal of Dairy Science 91, 847–856. https://doi.org/10.3168/jds.2007-0536

DeJarnette, J.M., Leach, M.A., Nebel, R.L., Marshall, C.E., McCleary, C.R., Moreno, J.F., 2011. Effects of sex-sorting and sperm dosage on conception rates of Holstein heifers: Is comparable fertility of sex-sorted and conventional semen plausible? Journal of Dairy Science 94, 3477–3483. https://doi.org/10.3168/jds.2011-4214

Elite Magazin, 2009. Optimierung auf hohem Niveau [WWW Document]. Elite Magazin. URL https://www.elite-magazin.de/dossiers/optimierung-auf-hohem-niveau-9321742.html (accessed 7.17.19).

Ettema, J.F., Thomasen, J.R., Hjortø, L., Kargo, M., Østergaard, S., Sørensen, A.C., 2017. Economic opportunities for using sexed semen and semen of beef bulls in dairy herds. Journal of Dairy Science 100, 4161–4171. https://doi.org/10.3168/JDS.2016-11333

GAMS Development Corporation, 2019. General Algebraic Modeling System (GAMS) Release 27.1.0.

Healy, A.A., House, J.K., Thomson, P.C., 2013. Artificial insemination field data on the use of sexed and conventional semen in nulliparous Holstein heifers. J. Dairy Sci. 96, 1905–1914. https://doi.org/10.3168/jds.2012-5465

Hlavac, M., 2018. Stargazer: well-formatted regression and summary statistics tables. Bratislava, Slovakia.

Holden, S.A., Butler, S.T., 2018. Review: Applications and benefits of sexed semen in dairy and beef herds. Animal 12, s97–s103. https://doi.org/10.1017/S1751731118000721

KTBL, 2020. Standard gross margins [WWW Document]. URL https://daten.ktbl.de/sdb/welcome.do (accessed 5.16.21).

KTBL, 2018. Betriebsplanung Landwirtschaft 2018/19 Daten für die Betriebsplanung in der Landwirtschaft. Kuratorium für Technik und Bauwesen in der Landwirtschaft, Darmstadt.

Kuhn, M.T., Hutchison, J.L., Wiggans, G.R., 2006. Characterization of Holstein heifer fertility in the United States. J. Dairy Sci. 89, 4907–4920. https://doi.org/10.3168/jds.S0022-0302(06)72541-3

Kuhn, T., Schäfer, D., Holm-Müller, K., Britz, W., 2019. On-farm compliance costs with the EU-Nitrates Directive: A modelling approach for specialized livestock production in northwest Germany. Agricultural Systems 173, 233–243. https://doi.org/10.1016/j.agsy.2019.02.017

Lengers, B., Britz, W., Holm-Müller, K., 2014. What Drives Marginal Abatement Costs of Greenhouse Gases on Dairy Farms? A Meta-modelling Approach. Journal of Agricultural Economics 65, 579–599. https://doi.org/10.1111/1477-9552.12057

Lengers, B., Britz, W., Holm-Muller, K., 2013. Comparison of GHG-Emission Indicators for Dairy Farms with Respect to Induced Abatement Costs, Accuracy, and Feasibility. Applied Economic Perspectives and Policy 35, 451–475. https://doi.org/10.1093/aepp/ppt013

LfL Bayern, 2018. LfL Deckungsbeiträge und Kalkulationsdaten [WWW Document]. URL https://www.stmelf.bayern.de/idb/default.html (accessed 11.26.18).

LfL Bayern, 2016. Zifo2 - Zielwert Futteroptimierung. LfL Bayern, Poing.

Maicas, C., Holden, S.A., Drake, E., Cromie, A.R., Lonergan, P., Butler, S.T., 2020. Fertility of frozen sex-sorted sperm at 4 × 106 sperm per dose in lactating dairy cows in seasonal-calving pasture-based herds. Journal of Dairy Science 103, 929–939. https://doi.org/10.3168/jds.2019-17131

Markowitz, H., 1952. Portfolio Selection. The Journal of Finance 7, 77–91. https://doi.org/10.2307/2975974

Martens, H., 2016. Leistung und Gesundheit von Milchkühen: Bedeutung von Genetik (Ursache) und Management (Wirkung). Tierarztl Prax Ausg G 44, 253–258. https://doi.org/10.15653/TPG-160312

McCullock, K., Hoag, D.L.K., Parsons, J., Lacy, M., Seidel, G.E., Wailes, W., 2013. Factors affecting economics of using sexed semen in dairy cattle. Journal of Dairy Science 96, 6366–6377. https://doi.org/10.3168/JDS.2013-6672

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics 21, 239–245. https://doi.org/10.2307/1268522

Murphy, C., Shalloo, L., Hutchinson, I.A., Butler, S.T., 2016. Expanding the dairy herd in pasture-based systems: The role of sexed semen within alternative breeding strategies. Journal of Dairy Science 99, 6680–6692. https://doi.org/10.3168/JDS.2015-10378

Norman, H.D., Hutchison, J.L., Miller, R.H., 2010. Use of sexed semen and its effect on conception rate, calf sex, dystocia, and stillbirth of Holsteins in the United States. J. Dairy Sci. 93, 3880–3890. https://doi.org/10.3168/jds.2009-2781

Østergaard, S., Chagunda, M.G.G., Friggens, N.C., Bennedsgaard, T.W., Klaas, I.C., 2005. A stochastic model simulating pathogen-specific mastitis control in a dairy herd. J. Dairy Sci. 88, 4243–4257. https://doi.org/10.3168/jds.S0022-0302(05)73111-8

Pedersen, L.D., Sørensen, A.C., Henryon, M., Ansari-Mahyari, S., Berg, P., 2009. ADAM: A computer program to simulate selective breeding schemes for animals. Livestock Science 121, 343–344. https://doi.org/10.1016/j.livsci.2008.06.028

R Core Team, 2017. R: A language and environment for statistical computing. Vienna, Austria.

Römer, A., Blum, E., Böttcher, I., Sanftleben, P., 2009. Lebensleistung und Nutzungsdauer von Milchkühen aus der Sicht einer effizienten Milchproduktion. Züchtungskunde 81, 341–360.

Römer, A., Boldt, A., Blum, E., 2018. Investigations on lifetime production and longevity of dairy cows, in: Landesforschungsanstalt für Landwirtschaft und Fischerei Mecklenburg Vorpommern (Ed.), Mitteilungen Der Landesforschungsanstalt Für Landwirtschaft Und Fischerei Mecklenburg-Vorpommern : Heft 60. Landesforschungsanstalt für Landwirtschaft und Fischerei Mecklenburg Vorpommern, Gülzow.

Römer, A., Harms, J., Volkmann, N., Blum, E., Flor, J., 2013. Verbesserung der Fruchtbarkeitssituation in hochleistenden Milchviehherden (Abschlussbericht No. 2/51). Landesforschungsanstalt für Landwirtschaft und Fischerei - Institut für Tierproduktion, Gülzow.

Schneichel, R., 2017. Weiblich auf Wunsch – wie gut funktioniert gesextes Sperma? veterinär spiegel 27, 119–121. https://doi.org/10.1055/s-0043-116448

Seidel, G.E., 2014. Update on sexed semen technology in cattle. Animal 8 Suppl 1, 160–164. https://doi.org/10.1017/S1751731114000202

Statistisches Bundesamt, 2018. Fachserie 3, Reihe 4.1, Viehbestand. Wiesbaden.

Thomsen, J., 2016. Lohnt sich der Einsatz für Rinderhalter? Bauernblatt 17.09.2016.

Weiske, A., Vabitsch, A., Olesen, J.E., Schelde, K., Michel, J., Friedrich, R., Kaltschmitt, M., 2006. Mitigation of greenhouse gas emissions in European conventional and organic dairy farming. Agriculture, Ecosystems & Environment, Mitigation of Greenhouse Gas Emissions from Livestock Production 112, 221–232. https://doi.org/10.1016/j.agee.2005.08.023

Wolfová, M., Wolf, J., Kvapilík, J., Kica, J., 2007. Selection for Profit in Cattle: II. Economic Weights for Dairy and Beef Sires in Crossbreeding Systems. Journal of Dairy Science 90, 2456–2467. https://doi.org/10.3168/jds.2006-615

Chapter 3
'Fruchtfolge': A crop rotation decision support system for optimizing cropping choices with big data and spatially explicit modeling[1]

Abstract

Deciding on which crop to plant on a field and how to fertilize it has become increasingly complex as volatile markets, location factors as well as policy restrictions need to be considered simultaneously. To assist farmers in this process, we develop the web-based, open source decision support system 'Fruchtfolge' (German for 'crop rotation'). It provides decision makers with a crop and management recommendation for each field based on the solution of a single farm optimization model. The optimization model accounts for field specific location factors, labor endowments, field-to-farm distances and policy restrictions such as measures linked to the EU Nitrates Directives and the Greening of the EU Common Agricultural Policy. 'Fruchtfolge' is user-friendly by automatically including big data related to farm, location and management characteristics and providing instant feedback on alternative management choices. This way, creating a first optimal cropping

---

[1] This chapter is published in the journal *Computers and Electronics in Agriculture* as:

Pahmeyer, C., Kuhn, T., Britz, W., 2021. 'Fruchtfolge': A crop rotation decision support system for optimizing cropping choices with big data and spatially explicit modeling. Computers and Electronics in Agriculture 181, 105948. https://doi.org/10.1016/j.compag.2020.105948

plan generally requires less than five minutes. We apply the decision support system to a German case study farm which manages fields outside and inside a nitrate sensitive area. In the year 2021, revised fertilization regulations come in force in Germany, which amongst others lowers maximal allowed nitrogen applications relative to crop nutrient needs in nitrate sensitive areas. The regulations provoke profit losses of up to 15% for the former optimal crop rotation. The optimal adaptation strategy proposed by 'Fruchfolge' diminishes this loss to 10%. The reduction in profit loss clearly underlines the benefits of our support tool to take optimal cropping decisions in a complex environment. Future research should identify barriers of farmers to apply decision support systems and upon availability, integrate more detailed crop and field specific sensor data.

Keywords

big data, Decision Support System, Nitrates Directive, Fertilization Ordinance, farm level simulation model

3.1 Introduction

Every year, farmers need to decide anew which crops to plant on each of their fields. Their choices need to reflect a growing number of determinants. On the individual field level, location factors such as soil types and crop rotational effects, as well as technological, structural, and economical factors need to be considered (Kuhlmann, 2015). At farm scale, the cropping plan needs to fit to the farmer's labor and machinery endowments. Furthermore, command-and-control measures related to agri-environmental legislation need to be considered. The German implementation of the EU Nitrates Directive as the core regulation to protect water bodies from nitrate emissions from agriculture provides a striking example. It prescribes complex field specific management standards, for instance depending on the chosen crop, its yields, and the nitrogen content of the soil.

In the past, multiple attempts at assisting decision makers with the 'cropping choice problem' have been made. Methodologically, mathematical programming (including linear programming) has proven to be a powerful tool for the analysis of resource allocation choices (Hazell and Norton, 1986). McCarl et al. (1977) used linear programming to create an income maximizing cropping pattern for commercial grain farms in the Midwest. Their approach required farmers to fill an input form with their data, subsequently being evaluated by researchers. However, without research extension interaction, farm planning use of the model was found to be not generally practical.

Subsequent approaches focused on extensions to linear programming models as the inclusion of risk modeling (Mußhoff and Hirschauer, 2004, 2006a), or applications in the context of policy analysis (Galán-Martín et al., 2015; Louhichi et al., 2010). However, all the models solely returned optimal crop shares at farm scale. Compared to optimizing the spatially explicit crop allocation, this significantly reduces data needs and model complexity but disregards the heterogeneity of the individual fields and their spatial characteristics. It eventually leads to a sub-optimal solution to the original planning problem and, when used as DSS, leaves the decision taker with the daunting task to allocate the proposed optimal shares at farm scale to individual fields. Only Radulescu and Radulescu (2012) describe a DSS based on a portfolio selection model for crop planning under risk, that provides the user with a crop recommendation on a per field basis. However, their approach requires manual input for all crop and field related data and does not incorporate policy restrictions and manure allocation.

Despite these efforts, models to support cropping choices based on mathematical programming have rarely been adopted by farmers and farm advisers (Mußhoff and Hirschauer, 2016). As one of the main reasons for the relatively low uptake of such models, referred to as decision support systems (DSS) when focused on supporting farmers' management choices, Mußhoff and Hirschauer (2016) identify the high data requirements of mathematical programming.

In the underlying manuscript, we present the web-based DSS 'Fruchtfolge' (German for crop rotation) which supports farmers' in making optimal crop and crop management choices in a complex environment[2]. Fruchtfolge provides its users with a crop recommendation and manure application strategy for each of their fields, automatically incorporating big data from multiple sources related to farm, location, and management characteristics. By combining these datasets, a highly detailed single farm model is created and solved in real-time in the background, without requiring extensive user input. The model automatically adheres to legal restrictions from the German Fertilization Ordinance (FO), implementing the Nitrates Directive, and the Greening obligations of the EU Common Agricultural Policy. Following best practices of 'user-centered design' (Parker and Sinclair, 2001; Rose et al., 2017, 2016), the maximum required time to create an initial optimal cropping plan is targeted at 5 minutes, including application signup and data entry.

The contribution of the paper is twofold. First, we present Fruchtfolge as an innovative and unique DSS targeting (German) farmers and farm advisors. Second, we apply it to an exemplary farm which faces tighter measures of the FO, mainly coming in force from 2021 onwards, in order to illustrates the benefits of Fruchtfolge to find optimal cropping plans in complex environments.

3.2 Materials and Methods

3.2.1 *Overview of the decision support system "Fruchtfolge"*

"Fruchtfolge" is built in an effort to create a user-centered, simple to use DSS to provide profit maximal field specific cropping choices and fertilization strategies. Its development is based on best practices in agricultural DSS design outlined by Rose (2016), and experiences from established DSS such as ValorE (Acutis et al., 2014) or vite.net® (Rossi et

[2] The Fruchtfolge DSS is hosted at the following URL: https://fruchtfolge.agp.uni-bonn.de/. Please see the supplementary video for a short overview of the Fruchtfolge DSS [English], as well as the user documentation hosted at https://fruchtfolge.agp.uni-bonn.de/documentation/ [German]

al., 2014). Emphasis is put on the DSS core factors ‘performance’ and ‘ease-of-use’.

Figure 1 displays a systematic overview of Fruchtfolge. Three main steps are required in order to receive a first optimization result by the DSS. First, the user needs to initially sign-up on the website choosing a password, providing an E-Mail address as its user-id and the address of the farm premises. The address is required for the calculation of farm-to-field distances at a later stage. Like other web services, upon completion of the initial signup, users can later login again to the DSS using their E-Mail address and password and find all so far entered input and results. In a second step, users are asked to enter their so-called customer reference number (CRN, ZID number in Germany) which is available for every farm having applied for direct payments under the EU Common Agricultural Policy. Subsequently, the necessary data to optimize a cropping plan is downloaded automatically in the background and combined to a first version of the mixed integer linear programming (MILP) model without further action required from the user. Once this initial model is solved, the user is presented with the optimal cropping plan in a table and a map view with supporting graphs. In addition, a so-called fertilizing planning sheets as required by the German Fertilization Ordinance (FO) are generated. Next, the user can adjust input parameters such as prices, costs, yields, or crop share constraints and re-run the model. In the following, the technical procedure of the data acquisition is further explained.

3.2.2 *Graphical User Interface and technical implementation*

The Graphical User Interface (GUI) of the Fruchtfolge DSS enables the communication between the user, the data base and the underlying bio-economic model. The GUI shields off details of the technical implementation from the user, allowing them to successfully use the DSS without requiring in-depth knowledge about the underlying model (Britz, 2014). As illustrated in the top part of Figure 1, the GUI is divided into three main parts: 1) The landing or login page, 2) data input pages (divided into sub-pages for fields, crops and constraints), and 3) the results page. Technically, the Fruchtfolge DSS is built as a progressive web application written in Node.js (server side) and JavaScript (client side). Opposed to traditional desktop applications which users have to download and install on their PC, progressive web applications are loaded on the fly and have the benefit of being portable across a whole range of devices (computers, tablets,

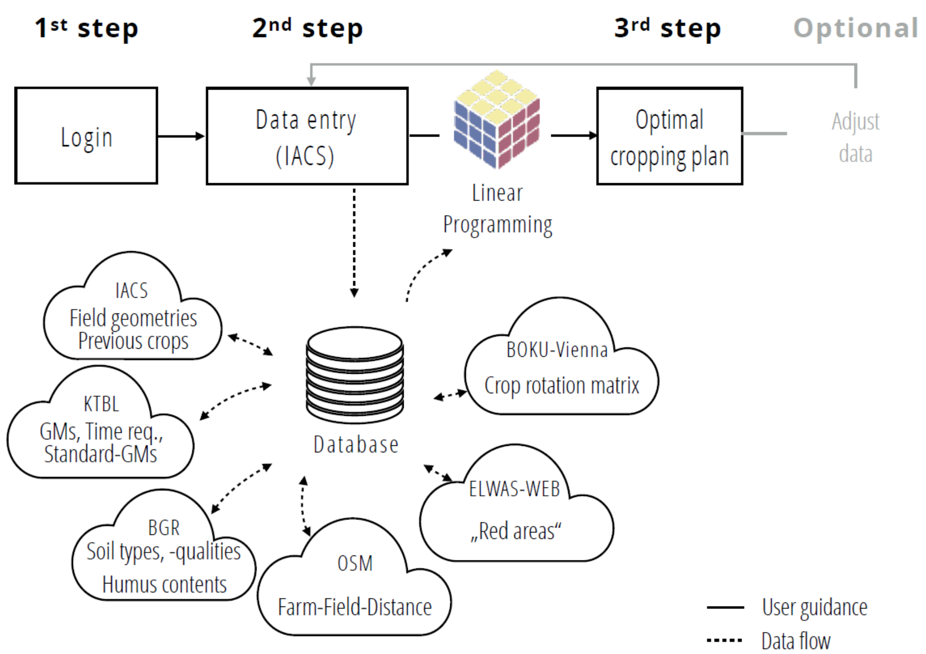


Figure 3.1 System architecture of Fruchtfolge. Further descriptions of the data sources can be found in Table 3.1

smartphones) and operating systems without requiring substantial changes to their codebase.

Users also automatically use the most up-to-date version of the DSS when visiting the website. A progressive web application requires constant internet access to deliver all of its features. According to a survey by the digital association Bitkom (2020), 82% of the German farmers already use digital technologies on their farm. Therefore, internet access and familiarity with digital technologies cannot be considered as a serious restriction regarding the use of a progressive web application. Fruchtfolge is open source and open access. Development of the application and its different sub-modules is steered from a public code versioning repository[3].

3.2.3 *Farm data import and big data use*

Detailed planning data is required for the optimization of a field specific farm cropping plan. In order to minimize manual data input, an importing routine in Fruchtfolge gathers automatically default information as detailed as possible for each field, crop and the farm as a whole. Users are free to overwrite each piece of information.

The different data sources automatically imported are displayed in Table 3.1 and in the bottom part of Figure 3.1. The CRN (ZID) provided by the user to gives access to the North Rhine-Westphalian IACS (Integrated Administration and Control System) database to collect data on the crops grown in previous years on each of the farm's fields along with their georeferenced locations and geometries. Jointly with the farmstead's address, this allows for calculating field specific farm-to-field distances based on Open Street Maps routing data. In addition, the import routines queries soil type, quality, and humus content for each field from the BGR database (BGR, 2016). Data from ELWAS-WEB (MULNV, 2020) locate fields in so called 'red' areas, i.e. nitrate sensitive zones with additional obligations according to the FO 2020 proposal.

[3] The main code versioning repository, as well as technical documentations of the different modules used in 'Fruchtfolge' can be found under: https://github.com/fruchtfolge

For each of the crops cultivated on the farm in the past, regionally specific (NUTS-2 level) yields, prices, and direct costs as elements of the Standard-Gross-Margin are gathered from the KTBL Standard-Gross-Margin database (KTBL, 2020) for the past 10 years, complemented by crop specific data on field operations from the KTBL database (KTBL, 2018).

Name	*Description*	*URL of the data*	*Application programming interface (API) to Fruchtfolge*
IACS database	The IACS (Integrated Administration and Control System) database includes field geometries as well as previous crops cultivated on the field for each farm in North Rhine-Westphalia.	https://www.lwk-verfahren.de/DownloadPortal/pages/index.action	https://github.com/fruchtfolge/elan-api
KTBL database	The KTBL (Kuratorium für Technik und Bauwesen in der Landwirtschaft) provides open data access to farm planning data such as regionalized historical yields, prices, and direct costs as well as field working operations depending on farm-field distances, soil types and field sizes.	https://srv.ktbl.de/doc/dev.en.html	https://github.com/fruchtfolge/KTBL-APIs

BGR maps	The BGR (Bundesamt für Geowissenschaften und Rohstoffe) provides maps regarding soil types, quality as well as humus contents.	https://www.bgr.bund.de/EN/Themen/Boden/boden_node_en.html	https://github.com/fruchtfolge/BGR-APIs
OSM	OSM (Open Street Maps) data is used to compute the field to farm distance for each field, relying on OSRM (Open Source Routing Machine).	https://github.com/Project-OSRM/osrm-backend	-
ELWAS-WEB	Outlines of 'red' areas according to the specification of Fertilization Ordinance at federal state level	https://www.elwasweb.nrw.de/elwas-web/index.jsf#	-
CropRota model, BOKU	The CropRota model (Schönhart et al., 2011) developed at the BOKU Vienna provides a value point matrix for different previous and subsequent crop combinations.	https://wpr.boku.ac.at/wpr_dp/DP-45-2009.pdf	-

Table 3.1 Source and description of external data used in the Fruchtfolge DSS

The KTBL database reports on time and machinery requirements, as well as variable and fixed costs depending on soil types, farm-to-field distances, yield levels and field sizes for individual field operations. The data is available for almost 100 crops, resulting in over 6,000,000 available data points. The time requirements for the single field operations allow estimating the required work time for a cropping plan in each month. Furthermore, the database also provides an estimate for the monthly available field working days depending on the field operation and region which can be interactively updated on demand. If the farmer enters available work hours per month, these data allow introducing monthly labor use constraints in the model. Basic parameters relating to cropping choice such as minimum rotational break years, previous crop effects (crop rotation matrix), and minimum soil requirements are taken from the CropRota model (Schönhart et al., 2011).

Nutrient contents, loss factors for the manure(s) and manure output per pig housing place are chosen according to the FO (BMEL, 2017). Along with the number of animal places provided by the user, this allows calculating the quantity of manure (liquid and solid) at farm level. The model depicts different nitrogen fertilizing levels and related yields for each crop based on N-response curves from Heyn & Olfs (2018). This is especially relevant under the FO 2020 where farmers have to reduce nitrogen fertilizer below the crop needs as in 'red' areas. Fruchtfolge either considers the restrictions of the FO 2017 or the FO 2020, depending on the farmer's choice. Primarily, both FO restrict the amount of manure and mineral fertilizer applied, as well as the legal time window of the application. As the regulations are part of the case study analysis, they are described in section 3.3.1.

The combination of the different data sources allows calculating gross margins and monthly labor requirements for each individual field and crop. For each field, the calculation reflects farm-to-field distance and size along with yield differences based on its soil quality and previous crop effects. The values are further differentiated for the following management options: varying levels of liquid and solid manure, cultivation of a catch crop (Boolean), manure application in autumn (Boolean), and different levels of nitrogen fertilizer reduction. Manure spreading options range from 0 m^3 ha^{-1} to 60 m^3 ha^{-1} in 5 m^3 steps reflecting typical manure barrel sizes.

Data on agronomic as well as on legislative constraints complements the information on farming operations and location characteristics. The field and crop specific minimum rotational breaks are complemented by maximum crop shares at farm level to avoid an overspecialization on the most profitable crops in the current year - the only one subject to optimization. To give an example, a minimum rotation break of two years for a crop on a field results in a maximum share of 33% ($1/(rotation\ break + 1)$) of the crop on the farms total cultivation area. Furthermore, the rules from the Greening obligation of the Common Agricultural Policy regarding minimum crop diversity and ecological focus area are considered in the DSS.

3.2.4 *Decision problem and optimization*

The calculations detailed above populate a matrix of all possible management options for each crop and field. All calculations are performed automatically in the background when new data are entered. An example of such a matrix is shown in Table 3.2. Each column of the matrix represents the (theoretically) possible cultivation options for one crop and field combination, characterized by the amount of manure to be spread, whether manure is applied in autumn, and whether a catch crop is cultivated before the main crop. If the FO 2020 proposal is active, an additional column indicates whether nitrogen fertilization should be reduced (and if yes, to which extent) for all fields that lie within a 'red' area as designated by the FO 2020.

This matrix depicts the decision space of the farmer. Without the support of the DSS, the decision maker would need to pick exactly one of these many options for each field, considering agronomic, economic, market, and legal constraints, partly at field, partly at farm level. Using a mathematical programming model, Fruchtfolge finds the optimal solution from the matrix which simultaneously considers all of these constraints. Based on its solution, Fruchtfolge proposes to the user (1) which crop to plant and (2) how much manure and mineral fertilizer to apply on each field.

Field	*Crop*	*Manure [m³/ha]*	*Autumn fertilization*	*Catch crop*
Field 1	Winter wheat	0	no	no
Field 1	Winter wheat	5	no	no
Field 1	Winter wheat	5	yes	no
…				
Field 20	Silage maize	60	yes	yes

Table 3.2 Example of a cropping matrix showing all possible cropping options for each field, crop, and manure combination.

As an example, a farm endowed with 20 fields and considering 5 different crops results in matrix with 5,200 columns, given 13 fixed manure spreading amounts, and the options of using or not manure application in autumn and a catch crop: $20 \cdot 5 \cdot 13 \cdot 2 \cdot 2 = 5.200$. If these 20 fields were located in a 'red' area, the matrix would even comprise 26,000 columns considering five possible N-reduction levels for each former option. Each column could either be chosen or not (Boolean), as we do not consider mixing crops or options on a field. This results in $2^{5,200}$ or even $2^{26,000}$ potential farm plans.

To address this complex decision problem for the user, a mixed integer linear programming model (MILP) is created and solved on the server side of the application. This offers a controlled technical environment with access to higher computing power, ensures that time for model generation and solve are independent from the user's hardware, and avoids installing the software for model generation and solution on the farmer's computer. As a first step, the matrix containing individual gross margins for each field, crop, manure amount, catch crop, and autumn fertilization option is created. Besides the gross margin, each column comprises entries which relate to farm-wide constraints: monthly labor needs, ecological focus area factors, as well as fertilizer demand. The resulting matrix enters a model which maximizes the farm's gross margin as the sum of the individual gross margins by field, crop, manure, catch crop, and autumn fertilization option multiplied with the (binary) decision variable indicating whether this option is active on the field or not. The model is written in the GAMS programming language (GAMS Development Corporation, 2019), and solved using the CPLEX MILP solver (IBM ILOG CPLEX, 2009). The source code of the model can be found in the supplementary material file[4].

For simplicity but without loss of generality, we summarized the manure amount, autumn fertilization, and catch crop options under the label *k*.

Following the notation used by Hazell and Norton (1986), the model can be written as follows:

$$\max_{v_{j,k,l}} tcm = \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} cm_{j,k,l} \cdot ha_l \cdot v_{j,k,l}$$

subject to

$$\sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} a_{r,j,k,l} \cdot ha_l \cdot v_{j,k,l} \leq b_r, \forall\, r = 1,2, \dots, R$$

where

tcm	Total expected contribution margin of the farm
$cm_{j,k,l}$	Expected contribution margin per ha for crop *j* combined with management option *k* on field *l*
ha_l	Size of field *l* in ha

[4] The source code of the model is available in the following versioning repository: https://github.com/fruchtfolge/model. The model version used in the manuscript can be found under the following DOI reference: 10.5281/zenodo.3626740.

$v_{j,k,l}$	Binary variable stating if crop j combined with management option k is present on field l
$a_{r,j,k,l}$	Coefficient of crop j combined with management option k on field l relating to resource or legal constraint r
b_r	Level of resource or legal constraint r

3.2.5 *Output*

Solving the model generally requires only a couple of seconds. Once the model is solved on the server, results are retrieved, processed and presented to the farmer in a sub page of the web application. An exemplary results page is displayed in Figure 3.2. The results page offers (1) a table showing the crop recommendation for each field, (2) a box indicating compliance with the greening legislation, (3) a pie chart with crop shares at farm level and information on the deviation from the optimal program when farmers adjust the cropping choice and management option (section 3.2.6), (4) two line charts, one displaying for the current year the monthly required work load and manure storage levels and a second one depicting profits over the last ten years under current year's plan at observed historic yields and prices, and finally (5) a map showing the spatial allocation of the different crops and the manure allocation. Furthermore, Fruchtfolge provides farmers with a field specific nitrogen and phosphate fertilizing planning sheet as required by the FO.

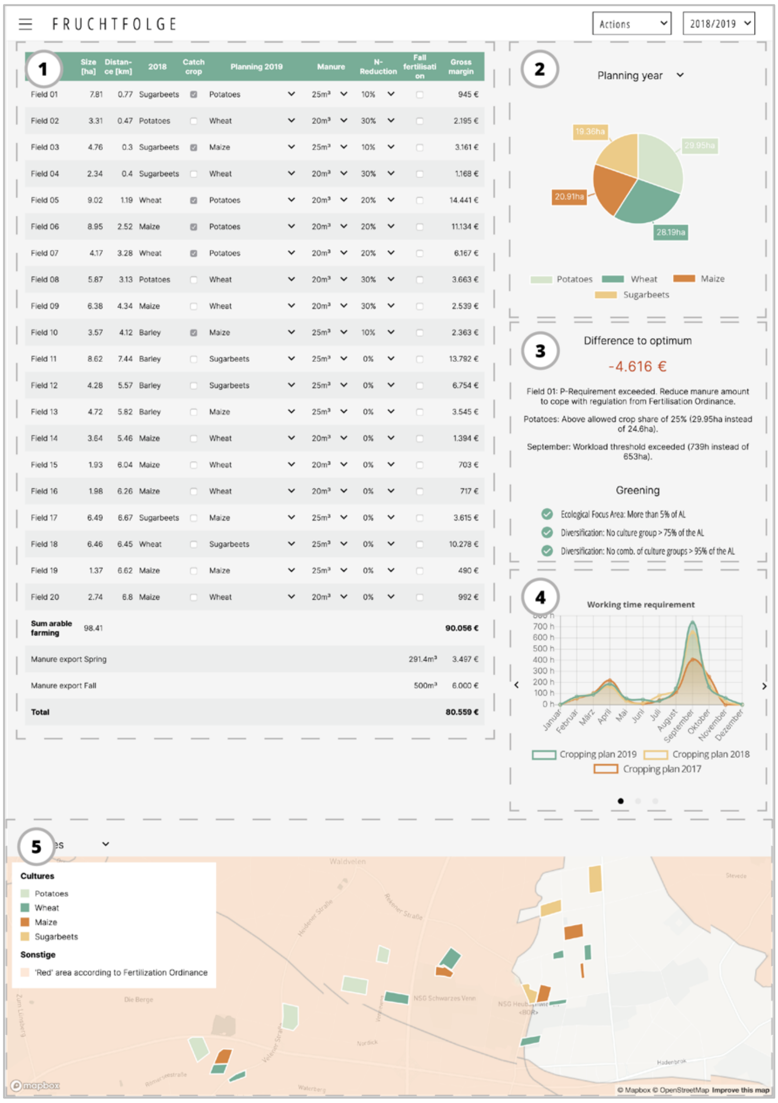


Figure 3.2 Exemplary results page of the Fruchtfolge DSS (translated).

The results page offers (1) a table showing the crop recommendation for each field, (2) a pie chart with crop shares at farm level, (3) a box indicating deviations against constraints and the influence of the violation on the farm profit, as well as compliance with the greening legislation when farmers adjust the recommended cropping choice and management, (4) line charts displaying the monthly required work load, monthly manure storage levels, as well as profits at observed historic yields and prices for the last ten years, and finally (5) a map showing the spatial allocation of the different crops and the manure allocation.

3.2.6 *Data adjustment and individualization*

Following the optimization, users are given two options of adjusting the optimization results. First, they may change the crop, manure application, catch crop, autumn fertilization, or N-reduction levels post simulation in the results page. When the user selects e.g. a different crop for a field, the results page is updated immediately, and a new info box is displayed. The info box will show the difference between the former optimized total contribution margin and the one reflecting the user's change. Note that an increase against the optimized total contribution margin is only possible if some constraint is violated. In that case, warnings show these violations, for instance exceeding maximal cropping shares or non-compliance with a measure from the FO. Hence, users can quickly perform 'what-if' scenarios and compare them with the optimization results. Providing the possibility of an ad-hoc sensitivity analysis aims to increase credibility in the DSS, and to reduce the black-box character of the underlying linear programming approach.

As a second adjustment option, users may alter the input data for the model. Opposed to the post simulation changes described before, changes to the input data are reflected in subsequent optimization runs. As previously stated, all the automatically acquired data can be changed. To give an example, users may add or remove fields, alter their geometries, change previous crops or mineralized nitrogen (N_{min}) contents. Regarding the crops, expected prices, yields, costs, maximum crop shares, labor requirements and previous crop effects can be adapted to the user's needs. In addition, fertilizing planning data such as target nitrogen amounts, manure nutrient contents, maximum manure application rates, and mineral fertilizer equivalents of manure can be changed.

3.3 **Case study**

In order to test the DSS and to illustrate its capabilities of finding the optimal crop and management choices in a complex environment, a hypothetical case study farm is generated. It is assumed to be located in the Borken region within the federal state of North Rhine-Westphalia, known for intensive livestock (mainly pig fattening) production (LWK NRW, 2014). The case study farm is assessed under varying policies, the FO 2017 as well as the FO 2020. The FO consists of numerous, partly interlinked measures which restrict the fertilizer management of farmers. The FO 2020 adds to the former version mainly additional restrictions in ‘red’ areas in which nitrate concentration targets are exceeded (see following section 3.3.1).

For the case study, three scenarios are modeled (see Table 3.3). The reference scenario (FULL-OPT-17) optimizes a cropping plan and fertilizing strategy under the FO 2017. It serves as a benchmark to calculate changes provoked by the FO 2020 as reflected in two additional scenarios. The first of these introduces the stricter obligations of the FO 2020 and evaluates their effect under the field specific cropping choice of the reference scenario. It is called FERT-OPT-20 as it only optimally adjusts the fertilization strategy to comply with new FO legislation but not the cropping plan. In the third scenario (FULL-OPT-20), Fruchtfolge finds the optimal cropping plan adaptation strategy under the proposed FO 2020 which minimizes compliance costs considering both changes in cropping choices and manure applications. All three scenarios use the same prices, yields as well as previous crops on each of the fields. Manure quantities not applied on the farm have to be exported and the related costs are added to the objective. Manure export costs of 12 € per m^3 are assumed (T. Kuhn et al., 2019b). The scenarios under FO 2020 are further differentiated by considering different shares of fields being situated in a red area.

Scenario	*Description*	*Fertilization Ordinance*	*Farmland in 'red' area*
FULL-OPT-17	Full optimization	2017	0%
FERT-OPT-20	Crop to field allocation fixed to reference scenario Fertilization strategy optimally adjusted to new Fertilization Ordinance	2020	0%, 50%, 100%
FULL-OPT-20	Full optimization	2020	0%, 50%, 100%

Table 3.3 Schematic overview of the scenario setup for the case study.

3.3.1 *German Fertilization Ordinance 2017 and 2020*

The FO implements the Nitrates Directive in Germany and was revised in 2017 after water quality benchmarks have been missed. The EU commission however sees the measures of the FO 2017 as insufficient to reach the environmental goals related to nitrate in ground and surface waters (Agra-Europe, 2019). Therefore, the FO has been anew revised in 2020 comprising distinct stricter measures (see Table 3.4).

	Fertilization Ordinance 2017	***Fertilization Ordinance 2020***	
	General changes	**General changes**	**Additional restrictions in 'red' areas**
Nutrient balance	Obligatory soil surface balance, surplus restricted	Nutrient balance abolished	-
Manure application	Limited to 170 kg N (nitrogen) ha^{-1} a^{-1}		Restriction applies at field instead of farm level
Fertilizing activities	Obligatory and predefined fertilizing planning based on N and P_2O_5 (phosphate) plant needs Only 10% of the autumn fertilization needs to be accounted for in the fertilizing planning calculation of the following year	Obligatory fertilizing planning and recording of every fertilizer application. Autumn fertilization has to be fully accounted for in the fertilizing planning calculation of the following year Minimum fertilizer efficiency coefficients for manure increased	-

Fertilizing restriction	Calculated plant need must not be exceeded	Minimum fertilizer efficiency coefficients for manure increased	Calculated plant need has to be undercut by 20%
Banning periods	Winter rape, winter barley, and catch crops can be fertilized in autumn with up to 60 kg N ha^{-1} a^{-1}	-	Winter rape, winter barley, and catch crops forbidden to fertilize in autumn
Catch crops	-	-	Obligatory catch crop cultivation for allowance of fertilizer application to following summer crops

Table 3.4 Overview on core changes from Fertilization Ordinance 2017 (BMEL, 2017) to 2020 (BMEL, 2020).

The FO consists of numerous, partly interlinked measures. Most changes from the FO 2017 to the FO 2020 are linked to so-called ‘red’ areas, which describe areas above groundwater bodies exceeding the target nitrate concentration or showing increasing trends. Already under the Fertilization Ordinance 2017, farmers had to fulfil additional measures in ‘red’ areas which were however little restrictive and not relevant for the assessed decision problem (see Kuhn (2017) for detailed description of Fertilization Ordinance 2017).

In the FO 2020, a prescribed and detailed fertilizing planning approach plays a major role and replaces former restrictions on nutrient surpluses. The fertilizer application, covered by manure or mineral fertilizer, is constrained based on each crop's need after subtracting different nutrient sources such as spring mineralization. The plant need is lowered by 20% for fields in 'red' areas, resulting in reduced fertilizer application. This reduction however applies in average on the affected fields, only, allowing for complex shifting between crops. Furthermore, the application of manure is restricted to 170 kg N (ha^{-1} a^{-1}, a threshold calculated at farm average under the FO 2017 and 2020. In 'red' areas, however, the threshold has to be met at field level. In addition, the mineral fertilizer equivalents of manure are increased in the fertilizing planning in the FO 2020. Finally, nitrate leaching in autumn should be reduced by the banning of fertilizer application to rape seed, winter barley and catch crops in spring as well as the obligatory catch crop cultivation before summer crops in the 'red' areas.

The measures of the FO 2020 render decisions on cropping choices and fertilizing more complicated. The described thresholds are added to the optimization process described in section 3.2.3. Also, the temporal limitations of fertilizer application are introduced as additional restrictions, returning the farm's optimal gross margin when meeting the requirements of the FO. The DSS thereby addresses the decision farmers have to take in the light of the stricter regulations of the FO 2020 such as (1) the adaption of cropping choice and fertilizer allocation inside and outside read areas, taking into account that N yield responses differ between crops, (2) the change of manure allocation on farm and manure export due to stricter application thresholds and banning of application in autumn, and (3) the economic assessment of summer crops due to costly obligatory catch crop cultivation.

3.3.2 *Case study farm characteristics*

To assess impacts of this spatially differentiated fertilizing restriction, the case study farm is given arbitrary 10 fields inside and 10 fields outside of a nitrate sensitive 'red' area in our medium scenarios. While the 20 fields are chosen arbitrarily for this case study, their shapes, previous crops, soil type, and quality correspond to actual fields. In total, the farm cultivates approx. 100 ha with an average field size of 5.6 ha outside and of 4.2 ha inside of the 'red' areas boundaries. Due to the shape of the 'red' area, the average field-to-farm distance is only 2 km for the fields inside and 6 km outside of it. Soil qualities and types are rather homogenous among the fields, with an average soil quality rating (Mueller et al. (2014)) of 64 inside and 57 outside of the 'red' areas boundaries. As discussed in section 3.2.3, regional crop yields, prices and direct costs were obtained from the KTBL-SGM database (KTBL, 2020). In the underlying assessment, the 10-year average of these values is considered as expected values for the planning period. N_{min} values for the fields are obtained from the North Rhine-Westphalian chamber of agriculture (LWK NRW, 2020a). To prohibit the generation of cropping plans that exceed the available labor endowment, peak labor constraints based on the previous year's cropping plans are introduced.

As previously stated, the case study farm is located in the livestock intensive region of Borken, Germany. The case study farm is given 2,000 pig fattening places which reflects the average in the region (Kreis Borken, 2020). It is assumed that the pigs are fattened using feed with reduced nitrogen and phosphate content (LWK NRW, 2018). Furthermore, a higher mineral fertilizer equivalent of 72% of the pig manure (compared to 60% stated in the FO) is assumed for the fertilizing planning, as suggested by planning data from the North Rhine-Westphalian chamber of agriculture (LWK NRW, 2020b). Due to the longstanding manure use in the area, N-target values are adapted accordingly (agrarheute, 2015).

The scenarios presented in Table 3.3 present different assumption on where fields are located: First, it is assumed that none of the fields are within a 'red' area. Second, half of the fields are in- and outside of the 'red' areas boundaries (reflecting the currently proposed boundaries). Third, all of the case study farms fields are simulated to be in a 'red' area. While the concept of 'red' areas has already existed in the FO 2017, the additional measures only included slightly stricter measures (BMEL, 2017). The varying size of the red area is therefore not included for the scenario FULL-OPT-17.

3.3.3 *Case study results*

For the given case study, we find profit losses induced by the stricter measures of the FO 2020 to vary largely depending on how many fields are situated in the 'red area'. The losses range between 1,500 € and 13,650 € for the given case study farm. By following Fruchtfolge's recommended cropping choices and fertilization strategies, the losses can be reduced by up to 4,700 € (see Figure 3.3).

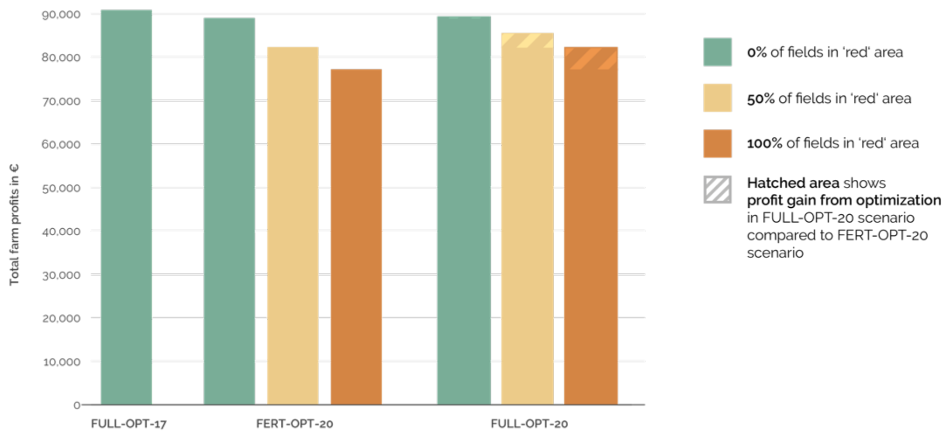


Figure 3.3 Total farm profits of the different scenarios depending on the share of fields in a 'red' area.

Under the FULL-OPT-17 scenario, the farm has a simulated profit of 90,506 € and faces manure export costs of 3,635 €. The farm grows maize on 44%, winter wheat on 18%, sugar beets on 20%, and potatoes on 17% of its land (Table 3.5). The farm cultivates about 26 ha of catch crops before seeding maize which allows to spread almost all available manure in autumn to avoid costly manure exports in this period.

In the FERT-OPT-20 scenarios, cropping choices are fixed to the results of FULL-OPT-17 scenario. This isolates the effects of the revised FO on the profit maximal fertilizing strategy and excludes the optimization of the cropping choices as a core feature of Fruchtfolge. The identical cropping plans of the FERT-OPT-20 and FULL-OPT-17 scenarios are displayed in the upper part of Figure 3.4. Under the FO 2020, instead of 10%, now 100%

of the nitrogen applied in autumn must be accounted for in the next year (see Table 3.4). Using catch crops to enable autumn fertilization of maize is no longer economically attractive. If no fields are located in a 'red' area, growing of catch crops is reduced to the point where it just fulfils the 5% minimum ecological focus area obligation under the Common Agricultural Policy. While this saves costs for catch crop cultivation, the manure not spread in autumn must be exported instead, leading to a net loss in profit of 1,591 € (-1.76% compared to the reference).

In the case of 50% of the case study farms fields being in a 'red' area, net profit loss increases to 8,850 € (-9.78% compared to the reference), driven mainly by two of the FO 2020 measures in 'red' areas. First catch crop cultivation is now mandatory before growing a summer crop such as maize and sugar beet. As maize and sugar beets may not be fertilized with manure in autumn under the FO 2020 obligations, manure exports in autumn increase. Second, the requirement to reduce the calculated plant need for nitrogen by an average of 20% reduces both crop yields and the total amount of manure which can be spread. While the yield loss leads to diminishing revenues, the reduced amount of manure that can be spread is additionally driving up manure export costs. These effects are amplified in the scenario where 100% of the case study farms fields are in a 'red' area: the net loss in profit is further increased to 13,658 € (-15.09% compared to the reference).

In the FULL-OPT-20 scenarios, cropping choices as well as manure allocation are optimized, illustrating the full potential of Fruchtfolge. In the simulation run where no fields are situated in a 'red' area, the farm can increase its profits by 188 € compared to the FERT-OPT-20 with no fields in the 'red' area. The profit increase is realized by an increase in the maize share at the expense of the wheat share (Table 3.5). As in the FERT-OPT-20 scenario with no fields in a 'red' area, autumn fertilization and related catch crop cultivation are completely abandoned. Giving up catch crop cultivation to a large degree and shifting manure application partly to spring frees labor in a peak period in autumn and allows for slightly increasing the maize share. While the expected gross margin for potatoes is higher than the one for maize, the freed labor allows for a higher return when the maize share is increased. This can be explained by the relatively high labor requirement of the potatoes, which (on average) require about 36 h/ha in autumn compared to 7.3 h/ha in maize.

In the FULL-OPT-20 simulation run where 50% of the fields are situated in a ‘red’ area, Fruchtfolge is able to increases the farms profits by 3,204 € compared to the respective simulation run in the FERT-OPT-20 scenario. The profit gain is realized by decreasing the maize share and expanding wheat and potato cultivation (see middle panel of Figure 3.4). These profit maximal adjustments reflect several interactions between crops due to labor constraints as well as subtle impacts of changes in the FO. This favors an expansion of wheat in ‘red’ areas as the farmer can apply more manure without exceeding application limits and avoid costly manure exports. Furthermore, in opposite to maize and potatoes as summer crops, wheat does not face costs of mandatory catch crop cultivation in a 'red' area of around 105 €/ha. Additionally, expanding the wheat share frees labor which can be used to increase potato cultivation.

In the FULL-OPT-20 simulation run where 100% of the fields are situated in a ‘red’ area, a profit increase of 4,710 € is realized by Fruchtfolge when compared to the same simulation run in the FERT-OPT-20 scenario. Similar to the simulation run with 50% of the fields in a ‘red’ area, the farm further decreases its maize share, and increases wheat and potato shares to their maximal shares at farm level (see bottom panel of Figure 3.4).

Scenarios	***Fertilisation Ordinance***	***Fields in 'red' area [%]***	***Maize***	***Wheat***	***Sugarbeet***	***Potatoes***
Reference / FERT-OPT	2017 / 2020	0%, 50%, 100%	44%	18%	20%	18%
FULL-OPT	2020	0%	45%	17%	20%	18%
FULL-OPT	2020	50%	30%	28%	20%	22%
FULL-OPT	2020	100%	24%	32%	20%	24%

Table 3.5 Optimal crop shares resulting from the optimization for the Reference and OPT scenario given different shares of fields in a ‘red’ area.

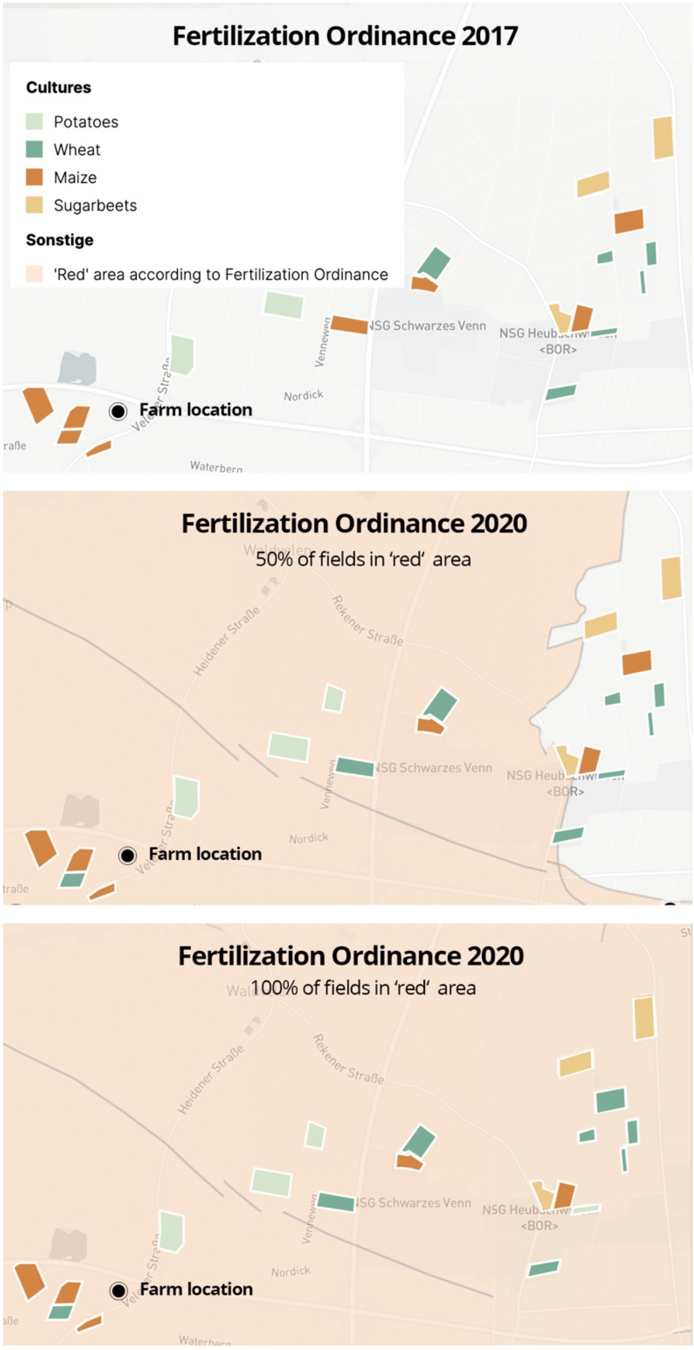


Figure 3.4 Optimized cropping plans as resulting from the Fruchtfolge DSS for the case study farm.

The top of the image is showing the optimal cropping plan under the FO 2017 (reference and FERT-OPT scenarios), the middle panel the optimal cropping plan under the FO 2020 with 50% of the fields in a 'red' area (FULL-OPT), and the bottom panel the optimized cropping plan with 100% of the fields in a red area (FULL-OPT).

3.4 Discussion

3.4.1 *Scope and technical implementation of Fruchtfolge*

The Fruchtfolge DSS supports farmers both with proposals for optimal cropping choices as well as fertilization strategies in accordance with the revised German FO. This renders Fruchtfolge not only useful for arable farmers, but also for livestock farmers optimizing their manure allocation while ensuring that certain shares of arable land are used for fodder production.

While the automatic farm data import was solely available for the state of North Rhine-Westphalia in Germany at the time of writing, future versions of the Fruchtfolge DSS will include automatic data import for all federal states of Germany, and possibly other countries as well. This feature is made publicly available by the *harmonie*[5] software package, which harmonizes farm subsidy application files across (federal) boundaries.

To the best of our knowledge, Fruchtfolge is the first cropping choice DSS to follow principles of user centered design outlined by Rose (2017, 2016), considering features of established DSS such as the sustainable vineyard management DSS vite.net® (Rossi et al., 2014).

This relates to fully automated data collection which only requires users to provide their CRN to access EU direct payment applications of their farm for an initial optimization. A user-friendly and visually attractive interface eases the communication between the DSS and the user, shielding it from details of the underlying economic programming model. Similar to vite.net®, Fruchtfolge aims at assisting the decision-maker by making recommendations which can be quickly explored with regard to alternatives and their consequences. Finally, as highlighted by vite.net®, providing the DSS as a web application enables continuous updates by the provider, and flexible access for decision makers. All these elements aim at overcoming the often-observed underuse of DSS at farm-scale.

[5] Hosted at https://github.com/fruchtfolge/harmonie

Fruchfolge captures a wide range of factors driving crop allocations to individual fields such as differences in gross margins, previous crop effects, minimal waiting times and restrictions related from command-and-control measures.

However, we deliberately do not expand Fruchfolge to cover diet optimization of animal herds and its interaction with optimal cropping choices. Far less automated data import is possible regarding the details of herd and, for instance, grass land management. Dynamics in livestock production and the inclusion of necessary intra-annual management options introduce numerous new aspects in the decision problem and require much more reflection on farm specifics.

Interactions with farmers and advisors revealed that constraints on minimal feed crop shares captured in Fruchtfolge are deemed as transparent and sufficient for a DSS with a focus on crop allocation and manure management.

Providing tools which help to understand why a certain solution of a larger programming model is economically optimal remains a challenge.

Fruchtfolge offers different views on the results (Figure 3.2) which also highlight interactions between cropping choices. It allows “challenging” the optimal solution and exploring consequences of alternative ones. Infeasible solutions, e.g. when choosing a crop and a field that would surpass a labor constraint for a certain month, are avoided by the introduction of slack variables with high penalties. Still, the optimal solution to the mixed integer problem underlying the Fruchtfolge DSS might remain a black box to some degree, which can undermine the trust of users in the DSS (Jakku et al., 2019). Further interviews with users can research this point and identify additional options for result analysis or automated support.

3.4.2 *Case study results*

The case study farm can increase its profits by 180 € up to 4,710 € by using the Fruchtfolge DSS when compared to an unchanged crop allocation reflecting the restrictions from the previous FO 2017. Increases in real world-cases are most likely considerably higher as users will also improve their crop and manure allocation at the benchmark (Mußhoff and Hirschauer, 2006b). The profit increases certainly outweigh the costs for the approx. 5 minutes of time required for a first optimal solution already tuned to farm specifics.

The announcements of the stricter measures outlined in the FO 2020 lead to nationwide protests from farmers in Germany, as reported in the media (Daily Mail, 2019). Due to Fruchtfolge's ability to outline a cost minimal compliance strategy for the FO 2020, the DSS may help to increase the acceptance of the stricter measures.

The case study highlights key drivers of farm-level impacts of the FO 2020. Compliance costs can be considerable and strongly depend on the share of farmland in 'red' areas. Therefore, our tool is shown to be of particular interest for farmers and farm advisers managing fields in such a 'red' area.

For the analysis, it was assumed that labor use should not exceed the labor use of the previous year's cropping plans. We found that the marginal profitability of an additional hour of labor in September could reach up to around 90 €. Case study results are hence rather sensitive to the available labor endowment in autumn. In interactive use, a farmer would probably allow for higher labor input in this period. This underlines the usefulness of interactive data updates and also points at new possibilities to exploit the dual solution as well to develop recommendations in future releases.

As previously stated, the assumed reduced N-requirements of the crops reflect long-standing manure fertilization present in the case study. This assumption dampens yield reductions and profit impacts of the required 20% reduction of nitrogen fertilizer application in a 'red' area. Given less efficient fertilizer management, optimization gains realized by Fruchtfolge will be higher. Note that all fertilizing parameters can be interactively adjusted in Fruchtfolge (section 3.2.6) in order to precisely reflect farm characteristics and famer's preferences.

Similar to the findings of Kuhn et al. (2019b), our results including compliance cost are quite sensitive to parameters related to fertilizer use. In this regard, literature finds strong efficiency differences in farm samples (LWK NRW, 2018; Osterburg and Techen, 2012) which can only partly be related to farm type and locational factors such as soil and climate. Compliance costs with the FO 2020 will therefore differ across farms as well as potential benefits from using Fruchtfolge. Both also depend on the assumed manure export costs as to some degree manure export is a central compliance strategy to the FO 2020. The costs chosen in the case study reflect current conditions for manure exports such as average transport distances. However, some cost increases are likely under the FO 2020 as many German livestock farms will need to expand exports, driving up transport distances and thus costs. Both the assumed high fertilizing efficiency in the case study and using current manure export cost render the reported compliance costs rather lower limits for actual ones in our case study farm. Again, the possibility to interactively change these assumptions in the DSS renders Fruchtfolge useful for evaluating possible impacts of higher manure exports costs on a particular farm.

3.4.3 *Implementation in practical use*

Musshoff and Hirschauer (2016) state that despite ongoing research efforts, mathematical programming methods have barely been adopted by farmers and farm advisers in Germany. As one of the main reasons, they argue that high data requirements impede the adoption of DSS using mathematical programming. Incorporating automation in data collection, following best practices of user centered design and lessons learned from established DSS, Fruchtfolge aims to overcome this implementation gap.

Our case study underlines the usefulness of applying a constrained optimization framework to determine which crop to grow on which field and how to fertilize it, especially in the light of a complex regulatory environment. Farmers may use “Fruchtfolge” to identify optimized production alternatives to their current production program which comply with the updated legislation and reflect manifold farm and field specific characteristic and restrictions. “Fruchtfolge” thus helps farmers to minimize compliance cost for the newest revision of the FO 2020. Also, Fruchtfolge helps farmers to avoid penalties due to accidental violations against legal frameworks, as the optimized cropping plan will automatically adhere to them and will warn farmers about violations.

Ongoing tests of the DSS with farmers are promising. Especially younger farmers (digital natives) show a high acceptance. Future research should evaluate the usefulness, design, and eventually the adoption of the Fruchtfolge DSS with farmers and lead the further development of the application. Multiple extensions to the current functionality are possible: future versions of the DSS could for example help farmers evaluate the profitability of agri-environmental measures on their farm, and thus improve the environmental footprint of the farm while increasing income. At present, the German agricultural administration digitizes reporting obligations and services for farmers. This process offers the chance to link DSS like Fruchtfolge to existing and widely-used digital platforms, and thereby promote the use of DSS in farming.

3.5 Conclusion

The Fruchtfolge DSS provides farmers with an economically optimal cropping and fertilizing plan without the need of time-consuming data input. In our case study, profit gains ranging from 180 € up to 4,710 € can be realized by using the DSS. Fruchtfolge reflects various legal constraints and thus helps farms to comply with the new FO in a cost minimal way. Due to its flexibility and design, farmers can easily carry out what-if scenarios and challenge the results of the underlying mathematical optimization model. This allows for “informed decisions” about alternative cropping and fertilizer management choices based on the economic, agronomic, and legal consequences compared to the optimized plan proposed by the DSS. Incorporating experiences from the literature about best practices in the design and implementation of a DSS, Fruchtfolge offers an attractive user interface and fast response times to overcome the “implementation gap” often prevalent with other DSS. With the increasing availability of site-specific sensor data, Fruchtfolge can be enhanced to incorporate even higher detailed farm specific data without requiring additional user interaction. Fruchtfolge is free and open source, and welcomes contributions to its codebase and documentation.

3.6 Acknowledgements

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2070 – 390732324.

3.7 References

Acutis, M., Alfieri, L., Giussani, A., Provolo, G., Guardo, A.D., Colombini, S., Bertoncini, G., Castelnuovo, M., Sali, G., Moschini, M., Sanna, M., Perego, A., Carozzi, M., Chiodini, M.E., Fumagalli, M., 2014. ValorE: An integrated and GIS-based decision support system for livestock manure management in the Lombardy region (northern Italy). Land Use Policy 41, 149–162. https://doi.org/10.1016/j.landusepol.2014.05.007

Agra-Europe, 2019. Kommission zwingt Deutschland zur Nachbesserung bei der Düngeverordnung. Agra-Europe 6, 35–36.

agrarheute, 2015. Stickstoffdüngung: So finden Sie die richtige Menge [WWW Document]. agrarheute. URL https://www.agrarheute.com/pflanze/stickstoffduengung-so-finden-richtige-menge-442024 (accessed 4.22.20).

BGR, 2016. Ackerbauliches Ertragspotential der Böden in Deutschland. Bundesanstalt für Geowissenschaften und Rohstoffe (BGR). [WWW Document]. URL http://www.bgr.bund.de/DE/Themen/Boden/Ressourcenbewertung-management/Ertragspotential/Ertragspotential_node.html (accessed 1.22.17).

Bitkom, 2020. Schon 8 von 10 Landwirten setzen auf digitale Technologien. Bitkom (e.V.). [WWW Document]. URL https://www.bitkom.org/Presse/Presseinformation/Schon-8-von-10-Landwirten-setzen-auf-digitale-Technologien (accessed 5.6.20).

BMEL, 2020. Verordnung zur Änderung der Düngeverordnung und anderer Vorschriften. Federal Ministry of Food and Agriculture (BMEL). [WWW Document]. URL https://www.bundesrat.de/SharedDocs/drucksachen/2020/0001-0100/98-20.pdf?__blob=publicationFile&v=2 (accessed 6.5.20).

BMEL, 2017. Verordnung zur Neuordnung der guten fachlichen Praxis beim Düngen. Federal Ministry of Food and Agriculture (BMEL) [WWW Document]. URL https://www.bundesrat.de/SharedDocs/drucksachen/2017/0101-0200/148-17.pdf?__blob=publicationFile&v=9 (accessed 5.6.20).

Britz, W., 2014. A New Graphical User Interface Generator for Economic Models and its Comparison to Existing Approaches [WWW Document]. German Journal of Agricultural Economics. https://doi.org/ISSN 0002-1121

Daily Mail, 2019. 10,000 farmers descend upon Berlin [WWW Document]. Mail Online. URL https://www.dailymail.co.uk/news/article-7727755/10-000-farmers-descend-Berlin.html (accessed 5.15.20).

Galán-Martín, Á., Pozo, C., Guillén-Gosálbez, G., Antón Vallejo, A., Jiménez Esteller, L., 2015. Multi-stage linear programming model for optimizing cropping plan decisions under the new Common Agricultural Policy. Land Use Policy 48, 515–524. https://doi.org/10.1016/j.landusepol.2015.06.022

GAMS Development Corporation, 2019. General Algebraic Modeling System (GAMS) Release 27.1.0.

Hazell, P.B.R., Norton, R.D., 1986. Mathematical programming for economic analysis in agriculture. Macmillan, New York.

Heyn, J., Olfs, H.-W., 2018. Wirkungen reduzierter N-Düngung auf Produktivität, Bodenfruchtbarkeit und N-Austragsgefährdung - Beurteilung anhand mehrjähriger Feldversuche, VDLUFA-Schriftenreihe. VDLUFA-Verlag, Darmstadt.

IBM ILOG CPLEX, 2009. V12. 1: User's manual for CPLEX. International Business Machines Corporation 46, 157.

Jakku, E., Taylor, B., Fleming, A., Mason, C., Fielke, S., Sounness, C., Thorburn, P., 2019. "If they don't tell us what they do with it, why would we trust them?" Trust, transparency and benefit-sharing in Smart Farming. NJAS - Wageningen Journal of Life Sciences 90–91, 100285. https://doi.org/10.1016/j.njas.2018.11.002

Kreis Borken, 2020. STATISTIKATLAS [WWW Document]. URL https://kreis-borken.de/fileadmin/kbor/FB_15/Statistik/Statistikatlas/atlas.html (accessed 2.4.20).

KTBL, 2020. Standard gross margins [WWW Document]. URL https://daten.ktbl.de/sdb/welcome.do (accessed 5.16.21).

KTBL, 2018. Betriebsplanung Landwirtschaft 2018/19 Daten für die Betriebsplanung in der Landwirtschaft. Kuratorium für Technik und Bauwesen in der Landwirtschaft, Darmstadt.

Kuhlmann, F., 2015. Landwirtschaftliche Standorttheorie: Landnutzung in Raum und Zeit, 1. Auflage. ed. DLG-Verlag, Frankfurt am Main.

Kuhn, T., 2017. The revision of the German Fertiliser Ordinance in 2017 (Discussion Paper No. 2), Discussion Paper. Universität Bonn, Bonn.

Kuhn, T., Schäfer, D., Holm-Müller, K., Britz, W., 2019. On-farm compliance costs with the EU-Nitrates Directive: A modelling approach for specialized livestock production in northwest Germany. Agricultural Systems 173, 233–243. https://doi.org/10.1016/j.agsy.2019.02.017

Louhichi, K., Kanellopoulos, A., Janssen, S., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Lansink, A.O., Ittersum, M.V., 2010. FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies. Agricultural Systems 103, 585–597. https://doi.org/10.1016/j.agsy.2010.06.006

LWK NRW, 2020a. Nmin-Richtwerte für Winterungen 2020. Landwirtschaftskammer Nordrhein-Westfalen (LWK NRW). [WWW Document]. URL https://www.landwirtschaftskammer.de/landwirtschaft/ackerbau/duengung/basisinfos/nmin-winterungen.htm (accessed 3.30.20).

LWK NRW, 2020b. Düngen mit Wirtschaftsdüngern. Landwirtschaftskammer Nordrhein-Westfalen (LWK NRW). [WWW Document]. URL https://www.landwirtschaftskammer.de/landwirtschaft/ackerbau/duengung/duengeverordnung/wirtschaftsduenger.htm (accessed 3.30.20).

LWK NRW, 2018. Nährstoffbericht 2017 - Wirtschaftsdünger und andere organische Dünger. Chamber of Agriculture North Rhine-Westphalia (LWK NRW) [WWW Document]. URL www.landwirtschaftskammer.de/landwirtschaft/ackerbau/pdf/naehrstoffbericht-2017.pdf (accessed 5.11.20).

LWK NRW, 2014. Landwirtschaft im Münsterland Daten - Fakten - Analysen. Landwirtschaftskammer Nordrhein-Westfalen (LWK NRW).

McCarl, B.A., Candler, W.V., Doster, D.H., Robbins, P.R., 1977. Experiences with farmer oriented linear programming for crop planning. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 25, 17–30. https://doi.org/10.1111/j.1744-7976.1977.tb02862.x

Mueller, L., Schindler, U., Shepherd, T.G., Ball, B.C., Smolentseva, E., Pachikin, K., Hu, C., Hennings, V., Sheudshen, A.K., Behrendt, A., Eulenstein, F., Dannowski, R., 2014. The Muencheberg Soil Quality Rating for Assessing the Quality of Global Farmland, in: Mueller, L., Saparov, A., Lischeid, G. (Eds.), Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia, Environmental Science and Engineering. Springer International Publishing, Cham, pp. 235–248. https://doi.org/10.1007/978-3-319-01017-5_13

MULNV, 2020. ELWAS Web [WWW Document]. URL https://www.elwasweb.nrw.de/elwas-web/index.jsf (accessed 2.4.20).

Mußhoff, O., Hirschauer, N., 2016. Modernes Agrarmanagement: betriebswirtschaftliche Analyse- und Planungsverfahren, 4., überarbeitete und erweiterte Auflage. ed. Verlag Franz Vahlen, München.

Mußhoff, O., Hirschauer, N., 2006a. Die Rehabilitation von Optimierungsverfahren? -Eine Analyse des Anbauverhaltens ausgewählter Brandenburger Markt- fruchtbetriebe. Schriften der

Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V. 495–507.

Mußhoff, O., Hirschauer, N., 2006b. Wie viel bringt eine verbesserte Produktionsprogrammplanung auf der Grundlage einer systematischen Auswertung empirischer Zeitreihen? Agrarwirtschaft 55, 175–187.

Mußhoff, O., Hirschauer, N., 2004. Optimierung unter Unsicherheit mit Hilfe stochastischer Simulation und Genetischer Algorithmen. Agrarwirtschaft 53, 264–279.

Osterburg, Techen, 2012. Evaluierung der Düngeverordnung – Ergebnisse und Optionen zur Weiterentwicklung (Abschlussbericht). National Working Group for the Evaluation of the Fertilization, Braunschweig.

Parker, C., Sinclair, M., 2001. User-centred design does make a difference. The case of decision support systems in crop production. Behaviour & Information Technology 20, 449–460. https://doi.org/10.1080/01449290110089570

Radulescu, C.Z., Radulescu, M., 2012. A Decision Support Tool Based on a Portfolio Selection Model for Crop Planning under Risk. Studies in Informatics and Control 21, 377–382. https://doi.org/10.24846/v21i4y201203

Rose, D.C., Parker, C., Fodey, J., Park, C., Sutherland, W.J., Dicks, L.V., 2017. Involving stakeholders in agricultural decision support systems: Improving user-centred design. International Journal of Agricultural Management 6, 10. https://doi.org/10.5836/ijam/2017-06-80

Rose, D.C., Sutherland, W.J., Parker, C., Lobley, M., Winter, M., Morris, C., Twining, S., Ffoulkes, C., Amano, T., Dicks, L.V., 2016. Decision support tools for agriculture: Towards effective design and delivery. Agricultural Systems 149, 165–174. https://doi.org/10.1016/j.agsy.2016.09.009

Rossi, V., Salinari, F., Poni, S., Caffi, T., Bettati, T., 2014. Addressing the implementation problem in agricultural decision support systems: the example of vite.net®. Computers and Electronics in Agriculture 100, 88–99. https://doi.org/10.1016/j.compag.2013.10.011

Schönhart, M., Schmid, E., Schneider, U.A., 2011. CropRota – A crop rotation model to support integrated land use assessments. European Journal of Agronomy 34, 263–277. https://doi.org/10.1016/j.eja.2011.02.004

Chapter 4
Data on a synthetic farm population of the German federal state of North Rhine-Westphalia[1]

Abstract

Farm-scale and agent-based models draw typically on detailed and preferably spatially explicit single farm data. Data protection standards however restrict or exclude their access, as for example in Germany. We provide data on a synthetic farm population of the German federal state of North Rhine-Westphalia, mainly based on the German Farm Structure Survey 2016 and plot specific crop data from 2019/2020. The population is derived from farm typology at administrative unit level to which the observed plots are allocated afterwards. The data contains 25,858 farms and covers 1.3 million ha of agricultural land, provided at plot scale in a geospatial vector and at farm scale in tabular format. For each plot, the managing farm (including the estimated farm's location), the number of livestock, the cultivated crop, as well as the corresponding administration units are indicated. Furthermore, spatial data such as yield information, soil characteristics, as well as monitoring data on environmental status are attached. The provided data allows for diverse analysis on the farm population in the federal state of North Rhine Westphalia with farm, agent-

---

[1] This chapter is published in the journal *Data in Brief* as:

Pahmeyer, C., Schäfer, D., Kuhn, T., Britz, W., 2021. Data on a synthetic farm population of the German federal state of North Rhine-Westphalia. Data in Brief 36, 107007. https://doi.org/10.1016/j.dib.2021.107007

based or different bio-physical models. Furthermore, it can serve as a test data set for models which require detailed and spatially explicit farm data.

Keywords

Synthetic farm population, farm typology, Germany, North Rhine-Westphalia, farm modelling, agent-based modelling

4.1 Data description

The data set provides a synthetic farm population with single farm data of the German federal state of North Rhine-Westphalia, derived by combining different secondary data sources. This is particularly useful for single farm and agent-based (ABM) models that often require spatially explicit and highly detailed single farm data. The resulting population covers 25,858 single farms and 1.3 million ha of agricultural land in the state. This corresponds to approximately 77% of all farms and 89% of all agricultural land. The state covers a diverse farm population, comprising approximately 8,600 specialized arable farms, 4,800 specialized pig farms, 9,500 specialized cattle farms, and 3,100 mixed farms, of varying sizes. They are distributed over different landscapes, such as fertile plains dominated by specialized arable farms, sandy plains with are large share of intensive animal production, and low-mountain ranges characterized by permanent grassland and cattle production. For every farm, estimations of its location, of its managed plots with observed crops and of its livestock numbers are provided. Single farm data at this level of detail is required for spatial explicit or population-wide analysis. However, it is usually not available in Germany due to data protection guidelines. If access is granted, publication of results is restricted, and the handling of the data is governed by complex rules. The synthetic population presented here provides an alternative which reflects key characteristics of the actual farm population without drawing on detailed single farm, data protected information. The provided farms, including their location, do not correspond to observed real-world farms. Instead, they reflect the distribution of key characteristics in the true population and correspond, in their entity, to observed statistical measurements. All underlying data sources are published and publicly available.

Different data sources are combined to derive the presented population (Table 4.3). The core sources are a farm typology from Kuhn and Schäfer (2018) and frequency tables of farms at commune level (LAU - Local Administrative Units), both based on the German Farm Structure Survey 2016. It is complemented by spatial explicit land use for the crop year 2019/2020, taken from the Integrated Administration and Control System (IACS) for the direct payments of the EU Common Agricultural policy. This land use data is linked to further spatial data such as yield information, soil characteristics, or monitoring data on environmental status.

The derived farm population is supplied at two scales and data formats at a Mendeley repository (DOI 10.17632/75wngh8x4j.1). First, single farm data for the population is provided in CSV format, with one row per farm. The variables, reported in the columns (Table 4.1), cover a unique farm ID, administrative units, longitude and latitude of the hypothetical farm location, livestock numbers, land use, information on plot size and plot-farmstead-distance, and a list of the managed plots. Second, data for each plot is provided in Shapefile format, reporting its exact spatial location as a polygon. The related attribute table contains additional information per plot (Table 4.2), covering among others plot and farm ID, plot size, cultivated crop, administrative units, soil parameters, environmental parameters, and regional crop yields. Linkage of the data set can draw on the unique farm ID provided for every plot in the shapefile, or the list of plot IDs reported for each farm in the CSV file.

Variable	Type	Description
farmId	Nominal	Unique ID for each farm
farmTypeCategory	Nominal	Groups farms according to their main farming activities which are defined based on the relative contribution of farming activities to the standard output following the EU typology (European Commission, 2008)
scr	Nominal	Soil-climate-region of the location of the farmstead
NAME	Nominal	LAU name of the location of the farmstead
LAU	Nominal	LAU code of the location of the farmstead
nuts3	Nominal	NUTS 3 code of the location of the farmstead
lng	Continuous	Longitude of the location of the farmstead
lat	Continuous	Latitude of the location of the farmstead
cows	Continuous	Number of cows [heads]
bulls	Continuous	Number of bulls[2] [heads]
pigs	Continuous	Number of pigs [heads]
sows	Continuous	Number of sows [heads]
farmSize	Continuous	Total farmland endowment [ha]
arableLand	Continuous	Arable land endowment [ha]
grassLand	Continuous	Permanent Grassland endowment [ha]
Wheats	Continuous	Cereal cultivation area [ha]

[2] Number of bulls calculated based on livestock units provided by Kuhn and Schäfer (2018).

Abbreviations: LAU - Local Administrative Units, NUTS - Nomenclature of Territorial Units for Statistics

RootCrops	Continuous	Root crops cultivation area [ha]
ArableFodder	Continuous	Arable fodder cultivation area [ha]
Oilseeds	Continuous	Oilseeds cultivation area [ha]
ProteinCrops	Continuous	Protein crops cultivation area [ha]
OrnamentalPlants	Continuous	Ornamental plants cultivation area [ha]
EnergyCrops	Continuous	Energy crops cultivation [ha]
avgPlotSize	Continuous	Average plot size of farm [ha]
medianPlotSize	Continuous	Median plot size of farm [ha]
deviationPlotSize	Continuous	Standard deviation of plot sizes of farm [ha]
avgPlotDistance	Continuous	Average plot-farmstead distance of farm [km]
medianPlotDistance	Continuous	Median plot-farmstead distance of farm [km]
deviationPlotDistance	Continuous	Deviation from average plot-farmstead distance of farm [km]
plots	Nominal	List of unique plot IDs assigned to the farm, separated by semi-colons

Table 4.1 Variables of data at farm-level (file *Farm_Population_NRW_farm_data*)

Variable	Type	Description
id	Nominal	Unique plot ID
farmId	Nominal	Unique farm ID of the farm managing the plot
plotSize	Continuous	Size [ha]
Distance	Continuous	Plot-farmstead distance [km]
FLIK	Nominal	Unique field block ID
cultivation	Nominal	Crop cultivated
NUTS1	Nominal	Federal state
NUTS2	Nominal	Administrative district

NUTS3	Nominal	County
LAU	Nominal	LAU code
NAME	Nominal	LAU name
N-content soil	Continuous	Total N in soil [g kg^{-1}]
P-content soil	Continuous	Total P in soil [g kg^{-1}]
K-content soil	Continuous	Extractable K in soil [g kg-1]
LC0_Desc	Nominal	Description of the most proximate location where N-/P-/K data was obtained from
Bodenzahl	Continuous	Soil value (German: Bodenwertzahl)
SQR	Continuous	Müncheberg soil quality rating
soilType	Nominal	Soil type
soilCode	Nominal	Soil type code
humus	Ordinal	Humus content [%]
redArea	Nominal	Plot in red zone (nitrate pollution hotspot according to DVO 2020)
eutrophicArea	Nominal	Plot in eutrophic area (phosphate pollution hotspot according to DVO 2020)
water_erosion_code	Discrete	Water erosion risk classes (0 = no risk, 1 = medium risk, 2 = high risk)
water_erosion_lvl	Ordinal	Water erosion risk classes in textual form (German)
wind_erosion	Nominal	High risk of wind erosion
slopeGradient	Ordinal	Slope gradient class [%]
WHG	Nominal	Affected according to §38a WHG (with distance indication)[1]
DVO	Nominal	Affected according to fertilization regulation §5 (with distance indication)[1]

slopeBufferMarginArea	Continuous	Area of the plot that is affected by either WHG and / or DVO [ha][3]
slopeBufferMarginPoly	-	A GeoJSON polygon outlining the part of the plot affected by the slope gradient scenery [1]
sugarbeets_yield	Continuous	Average sugar beet yield at NUTS 3 level [dt ha^{-1}]
oats_yield	Continuous	Average oat yield at NUTS 3 level [dt ha^{-1}]
summer_barley_yield	Continuous	Average summer barley yield at NUTS 3 level [dt ha^{-1}]
triticale_yield	Continuous	Average triticale yield at NUTS 3 level [dt ha^{-1}]
winter_wheat_yield	Continuous	Average winter wheat yield at NUTS 3 level [dt ha^{-1}]
potatoes_yield	Continuous	Average potato yield at NUTS 3 level [dt ha^{-1}]
corn_silage_yield	Continuous	Average corn silage yield at NUTS 3 level [dt ha^{-1}]
winter_rye_yield	Continuous	Average winter rye yield at NUTS 3 level [dt ha^{-1}]
rapeseed_yield	Continuous	Average rapeseed yield at NUTS 3 level [dt ha^{-1}]
winter_barley_yield	Continuous	Average winter barley yield at NUTS 3 level [dt ha^{-1}]

Table 4.2 Variables of data at plot-level (file Farm_Population_NRW_plot_data)

[3] Variable not presented if not applicable for the plot. Abbreviations: DVO – German Fertilization Ordinance, K - potassium, LAU - Local Administrative Units, N - nitrogen, NUTS - Nomenclature of Territorial Units for Statistics, P - phophorus, WHG - German Federal Water Act.

4.2 Materials and Methods

In the following, we document the methodology for creating the synthetic farm population for the German federal state of North-Rhine Westphalia. The population generation can be understood as descriptive research as we aim at characterizing and depicting the farm population in North Rhine-Westphalia without testing any hypothesis or drawing conclusions. It consists of three major steps as outlined in Figure 4.1. First, we generate farm frequency tables at LAU level. Second, we process the contingency tables sourced from the farm typology by Kuhn and Schäfer (2018), match it to a fitting farm from the previously generated frequency tables at LAU level, and create a spatially implicit farm population. Third, the farms are assigned random locations within the boundaries of their LAU. Finally, the observed plots, sourced from the publicly available IACS dataset, are assigned to the individual farms, based on their aspired farm size, crop cultivation specialization, and grassland endowment. The latter steps then turn the spatially implicit into a spatially explicit farm population. The created synthetic farm population is linked to further spatial data on yields, soil characteristics, and monitoring data on environmental status.

The data sources used in our methodology are outlined in Table 4.3. The creation of the synthetic population draws on two sources, the German Farm Structure Survey and the IACS data on land use. The Farm Structure Survey, carried out every three to four years, provides single farm data for all farms above a certain size threshold. It is the basis for the typology by Kuhn and Schäfer (2018) as well as the official farm statistics which are used to create the frequency tables. The typology contains only the most important farm types and, therefore, the estimated population does cover neither all farms nor all agricultural land in North Rhine-Westphalia (see Kuhn and Schäfer (2018) for details). IACS data on land use for each plot are reported annually by the farmers to determine direct payments from the EU Common Agricultural Policy. The agency collecting the data offers public access to them, however without information on the farmer managing the plot, and aggregating single crop information mostly to group of crops. As the data are spatially explicit, further spatial data relevant for agricultural land use can be attached as reported in Table 4.3. Note that except for the farm typology created by Kuhn and Schäfer (2018), most dataset are readily available in other German federal states as well.

Name	Description	Source of data
Sources for population creation		
Farm typology by Kuhn & Schäfer [1]	A farm typology for the German federal state of North Rhine-Westphalia based on the Farm Structure Survey 2016	http://www.ilr.uni-bonn.de/agpo/publ/dispap/download/dispap18_01.pdf
Integrated Administration and Control System (IACS/INVEKOS) data 2019/2020	Declared and verified eligible parcels for the EU funding, open data from the Chamber of Agriculture North Rhine-Westphalia	https://www.opengeodata.nrw.de/produkte/umwelt_klima/bodennutzung/landwirtschaft/
Official farm statistics from IT NRW 2018	Official farm statistics on LAU and NUTS 3 levels based on the Farm Structure Survey 2016	https://www.it.nrw/statistik/wirtschaft-und-umwelt/land-und-forstwirtschaft/struktur-der-landwirtschaftlichen-betriebe
Spatial data linked to plots		
Administrative location units (variables NUTS 1, NUTS 2, NUTS 3, LAU)	Spatial data on administrative location units at NUTS 1, NUTS 2, NUTS 3 and LAU level	https://github.com/eurostat/Nuts2json https://www.opengeodata.nrw.de/produkte/geobasis/vkg/dvg/dvg1/
Nutrient content soil (variables N-content soil, P-content soil, K-content soil)	Total N (defined using ISO 11261:1995 method), total P (defined using ISO 11263:1194 method), extractable K (defined using USDA−NRCS, 2004	https://esdac.jrc.ec.europa.eu/content/lucas2015-topsoil-data#tabs-0-description=1

	method) based on LUCAS 2015 TOPSOIL data	
Soil value (variable Bodenzahl)	German soil quality ranking between 0 and 100	https://www.geoportal.nrw/suche?lang=de&searchTerm=3E7CC528-6560-4BBE-AAB0-7DE2417EF993
Soil quality rating (variable SQR)	Müncheberg soil quality rating, indicator-based assessment of soil quality and crop yields potential described by Mueller et al. (2014)	https://www.bgr.bund.de/DE/Themen/Boden/Ressourcenbewertung/Ertragspotential/Ertragspotential_node.html
Soil type (variables soilType, soilCode)	Soil type and corresponding code based on the classification of the LUFA NRW	https://www.geoportal.nrw/suche?lang=de&searchTerm=3E7CC528-6560-4BBE-AAB0-7DE2417EF993
Humus content of soil (variable humus)	Median content of organic matter in the topsoil, given in percent classes (e.g. '1-<2%')	https://www.bgr.bund.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_Karten_Datenbanken/Themenkarten/HUMUS1000OB/humus1000ob_inhalt.html
Nitrate and phosphate pollution hot spots (variables redArea, eutrophicArea)	Area defined as nitrate and phosphate pollution hot spots in accordance with $13 DVO	https://www.opengeodata.nrw.de/produkte/umwelt_klima/wasser/duev/
Erosion risk (variables water_erosion_code, water_erosion_lvl, wind_erosion)	Risk classes for water erosion and areas at risk for wind erosion linked agri-environmental measures	https://www.opengeodata.nrw.de/produkte/umwelt_klima/bodennutzung/landwirtschaft/
Slope and distance to surface waters (variables slope gradient, WHG, DVO,	Sloped areas neighboring surface waters with management restrictions according to WHG §38a and DVO §5	https://www.opengeodata.nrw.de/produkte/umwelt_klima/bodennutzung/landwirtschaft/

slopeBufferMarginArea, slopeBufferMarginPoly)		
Regional yields	Regional yield level for main arable crops based on data from 2014-2017	https://flf.julius-kuehn.de/webdienste/webdienste-des-flf/ernteertraege.html

Table 4.3 Data sources

Abbreviations: DVO – German Fertilization Ordinance, K - potassium, LAU - Local Administrative Units, N – nitrogen, NUTS - Nomenclature of Territorial Units for Statistics, P – phosphorus, WHG – German Federal Water Act.

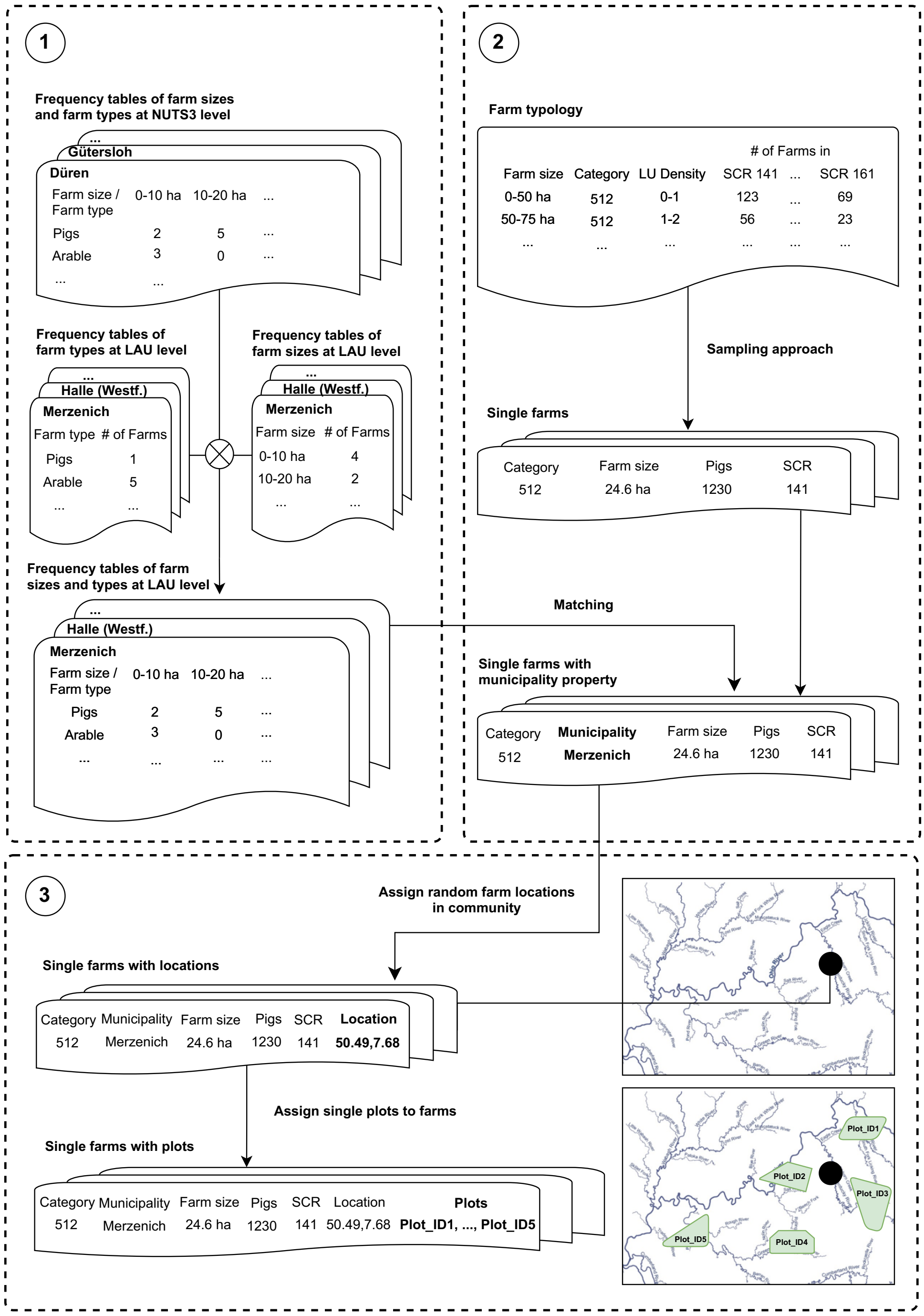


Figure 4.1 Overview on the synthetic farm population generation for the German federal state of North-Rhine Westphalia.

4.2.1 *Generation of frequency tables at municipality level*

This section presents the development of frequency tables which contain the frequencies of different farm types and size classes at the LAU level. The underlying code in the programming language GAMS is provided in a software versioning system[4]. Official statistics provide frequency tables for farm type and size classes, total utilized agricultural land, and agricultural land differentiated by land use at NUTS 3 and higher level, only. At LAU level, only vectors on the frequency of different farm types and size classes are reported. Our approach estimates probable frequency tables at LAU level, drawing on the frequency tables at NUTS 3 level and the vectors at LAU level. We use solely data from IT NRW 2018 for the estimation (Table 4.3).

The constraints of the estimation framework relate to adding up conditions at LAU level. Let $x_{s,c,t}$ denote the unknown number of farms of a certain size class s and type of specialization t in each of the 396 LAU. c, $d_{s,c}$ and $d_{t,c}$ are the given data on the number of farms of a certain size class, respectively, type, and d_c on the total number of farms. The following adding up conditions (eq. 4.1-4.3) should hold for any estimated frequency table of the farming population at LAU level $x_{s,c,t}$:

$$\sum_{s} x_{s,c,t} = d_{t,c} \tag{4.1}$$

$$\sum_{t} x_{s,c,t} = d_{s,c} \tag{4.2}$$

$$\sum_{s,t} x_{s,c,t} = d_{c} \tag{4.3}$$

After defining the adding up conditions on LAU level, we apply the same approach on NUTS 3 and 2 level as sometimes cells in frequency tables are left blank due to data protection rules. Let k denote county (NUTS 3, Kreis, 29 units) and r district (NUTS 2, Regierungsbezirk, 5 units) which are the

[4] The code is hosted at the following Github repository: https://github.com/MS-Dave/farms_nrw

two administrative units above LAU level where frequency tables on the number of farms by size class and type are available. Taking this additional information into account, we can add the following adding up conditions (eq. 4-5) from LAU to NUTS 3 and from NUTS 3 to NUTS 2 to the estimation framework.

$$\sum_{c \in k} x_{s,t,c} = d_{s,t,k} = x_{s,t,k} \tag{4.4}$$

$$\sum_{c \in r} x_{k,t,c} = d_{s,t,r} = x_{s,t,r} \tag{4.5}$$

The estimation problem is defined as a highest posterior density problem (HPD)[5]. We assume a-priori the relative, but unknown shares of the distribution by size class and type at LAU level are equal to the observed one at NUTS 3 level. Thus, s in the objective function (eq. 6) denotes the shares describing the empirical distribution observed at NUTS 3 level:

$$\min \sum_{s,t,c} \frac{x_{s,t,c}}{d_c} - s_{s,t,c} \tag{4.6}$$

The resulting estimates for x are real numbers and not counts, as required for the frequency tables. In order to convert them into a distribution of integers, we introduce bounds around each estimated $x_{s,t,c}$ representing its floor and ceiling values. We next construct a new estimator where equation 6 is replaced by an objective function which shifts the value towards an integer. To do that the estimator minimizes the squared difference between the estimates of $x_{s,t,c}$ and a number which is smaller than its lower bound if it is closer to the lower bound or higher than its upper bound otherwise. Any $x_{s,t,c}$ which is already an integer is automatically fixed as its floor and ceiling are identical. The additional estimation is repeated several times until all estimates x take on integer values

[5] For further discussion on HPD estimators please refer to Heckelei et al. (2008).

4.2.2 *Generation of the farm population*

Given the number of farms at LAU level in each category, we start to generate the spatially explicit farm population. The corresponding code, as well as the linkage of spatial data to the farm population is written in the programming language Node.js, and provided in a software versioning repository[6].

As stated previously, the analysis builds on the North-Rhine Westphalian farm typology published by Kuhn & Schäfer [1]. It differentiates farm types according to (1) type of farming, (2) size class in ha und (3) livestock density in livestock units (LU) per ha in different classes, and reports their numbers at the level of so-called soil-climate-regions (SCRs, generally consists of multiple NUTS 3 regions), which reflect zones of similar farming conditions.

The typology reports for each farm type the number of farms as well as statistics (mean, median, standard deviation) on core farm characteristics, including among others farm size in ha, arable and grassland endowment, and livestock density in LU, in total and for the animal categories pigs, sows, dairy cows, and other cattle. Based on the frequency tables, a sampling approach generates a matching farm population, and attaches characteristics according to the information found in the farm typology to each farm. The information on size and livestock density in the farm typology refers to certain classes (e.g. 0-50 ha, or 0-1 LU), accordingly, a truncated normal distribution is assumed when drawing these characteristics for a hypothetical farm in a cell of the typology.

In order to determine the nonparametric skew (S) of each farm characteristics, we calculate the skew using

$$S = \frac{\mu - \nu}{\sigma}$$

where μ is the populations mean, ν the populations median, and σ is the populations standard deviation for the given variable.

---

[6] The code is hosted at the following Github repository: https://github.com/chrispahm/farm-population-nrw

Due to data protection, statistics on farm characteristic for some farm types has been blackened in the farm typology. In these cases, average values given the farms size cluster are assumed for the variables with missing distribution data. For instance, farms in a size cluster between 0 ha and 50 ha would be assigned to 25 ha. Detailed comments in the relevant code sections report further assumptions made in the sampling approach. Following this sampling approach, a spatially implicit farm population containing a list of farms with specific values for each farm characteristic is generated (Figure 4.1).

So far, the location of a farm can only be assigned at SRC level. To advance here, the farm frequency tables at LAU level described in the section 4.2.1 are used. For each farm in the farm population, a farm at LAU level is matched , considering a) its farm type, b) its size class (e.g. 10-30 ha) and c) that the LAU falls in the SCR of the farm. Once a matching farm from the frequency tables is found, the LAU property from the match is added to the respective farm of the farm population. Thereby, the number of farms in the frequency table of this LAU that are not yet distributed is decreased. The LAU frequency tables and the farm typology are not completely harmonized, even if they stem originally from the same raw data source. Therefore, few farms cannot be matched, as each farm in the frequency table is at SCR level only used once in the matching procedure. In these cases, a farm is chosen for which solely the SCR and size cluster matches.

4.2.3 *Linking spatial data to the farm population*

Once each farm is assigned to a LAU, specific farm locations are designated to the farms of the generated farm population. Farm locations are assumed to be a random vertex from an arbitrarily chosen field (polygon) within the boundaries of the LAU the farm belongs to. If more than 50% of the farm's land endowment is arable land, only arable plots are considered for the farm location. Respectively, if more than 50% of the farms land endowment is grassland, only grassland plots are considered. If a vertex has already been defined as a farm location, the algorithm is recursively called until a new, unused farm location is found.

In a final stage of the generation of the spatially explicit farm population, individual plots are assigned to the farms. The plots in the federal state are shuffled to guarantee a random order, and afterwards evaluated for their suitability for a given farm. The assignment procedure generally differentiates between grassland and arable plots.

For grassland plots, farms that are within a 30 km driving radius around the plot are filtered and sorted by distance in ascending order. To maximize the efficiency of the algorithm, filtering of the farms is done using a spatial index based on a flat k-d tree as proposed by Bentley (1975). Subsequent to the farm filtering, the closest farm is searched for where the sum of the current plot and the current grassland endowment of the farm does not exceed the aspired grassland endowment of the farm (including a buffer of 5%). Also, a check is incorporated prohibiting a farm to exceed its farm size cluster. If a matching farm is found, the current plot is added to the farm, otherwise the plot is added to a list of unused plots.

For arable plots, farms within close proximity (5 km) are filtered and sorted based on their suitability for the given crop cultivated on the current plot. This is done to increase the probability of e.g. a specialized cereal producing farm to obtain plots cultivated with cereals. Again, a farm is searched for where the sum of the current plot and the current arable land endowment of the farm does not exceed the aspired arable land endowment of the farm (including a buffer of 5%). In addition to the check prohibiting the farm to exceed its farm size cluster, another check is incorporated prohibiting farms above 10 ha (the threshold where the EU Greening obligation becomes binding) to exceed certain crop shares. In case no suitable farm is found within 5 km for the given plot, the radius is increased to 30 km. Also, the sorting is solely based on the farm to field distance. If a matching farm is found, the current plot is added to the farm, otherwise the plot is added to the list of unused plots.

Farms that are within 95% of their aspired farm size are labelled as finished and removed from the list of farms considered in the evaluation of the following plots. After the first round of the assignment procedure, the list of unused plots is iterated over again.

In this second round, the plots are assigned to the most proximate farm as long as adding the plot to the farm does not exceed the farms overall aspired farm size. Here, the aspired arable and grassland properties are ignored, allowing for a deviation of these values in case no sufficient arable or grassland is available in the region of the farm.

Using the approach outlined in this section, matching the approx. 700.000 plots and 25.500 farms in the federal state of North Rhine-Westphalia takes less than 10 minutes.

4.3 Acknowledgements

This research has received funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - EXC 2070 – 390732324 and by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 817566). Part of this research was done within the Bioeconomy Science Center (BioSC) FocusLab "Transform2Bio". The scientific activities of the BioSC were financially supported by the Ministry of Culture and Science within the framework of the NRW Strategieprojekt BioSc (No. 313/323-400-002 13).

4.4 References

Bentley, J.L., 1975. Multidimensional binary search trees used for associative searching. Commun. ACM 18, 509–517. https://doi.org/10.1145/361002.361007

European Commission, 2008. Commission Regulation (EC) No 1242/2008 of 8 December 2008 establishing a Community typology for agricultural holdings, OJ L.

Heckelei, T., Mittelhammer, R.C., Jansson, T., 2008. A bayesian alternative to generalized cross entropy solutions for underdetermined econometric models (Discussion Paper No. 1548-2016–132440), Discussion Paper Series. Institute for Food and Resource Economics, Bonn. https://doi.org/10.22004/ag.econ.56973

Kuhn, T., Schäfer, D., 2018. A farm typology for North Rhine-Westphalia to assess agrienvironmental policies (Discussion Paper No. 1), Discussion Paper Series. Institute for Food and Resource Economics, Bonn.

Mueller, L., Schindler, U., Shepherd, T.G., Ball, B.C., Smolentseva, E., Pachikin, K., Hu, C., Hennings, V., Sheudshen, A.K., Behrendt, A., Eulenstein, F., Dannowski, R., 2014. The Muencheberg Soil Quality Rating for Assessing the Quality of Global Farmland, in: Mueller, L., Saparov, A., Lischeid, G. (Eds.), Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia, Environmental Science and Engineering. Springer International Publishing, Cham, pp. 235–248. https://doi.org/10.1007/978-3-319-01017-5_13

Chapter 5
Single plots or shares of land - How modeling of crop choices in bio-economic farm models influences simulation results[1]

Abstract

In bio-economic farm models, crop choices are generally depicted as shares of land types which are aggregates of plots with similar characteristics. The ongoing process of digitalization allows access to highly detailed, spatially explicit farm data and facilitates to represent single plots instead. In our paper, we examine how different approaches to model crop choices influence the results of an arable farm in a bio-economic model. Three possible approaches are considered: ‘single plots’ with one crop per season, crop shares of land differentiated by soil type, called ‘categorized’, and crop shares on all arable land, termed ‘aggregate’. The analysis is conducted using a highly detailed, spatially explicit dataset of 8,509 arable farms located in the German federal state of North Rhine-Westphalia. Our analysis indicates that the ‘aggregate’ and ‘categorized’ land endowment approaches produce similar simulation results, which however diverge from the ‘single plot’ approach. We find that on average, crop choices per farm differ by 11% between the spatially explicit ‘single plot’ and the ‘aggregate’ land

[1] This chapter is published in the “Food and Resource Economics, Discussion Paper” series as: Pahmeyer, C., Kuhn, T., Britz, W. (2021): Single plots or shares of land - How modeling of crop choices in bio-economic farm models influences simulation results. Discussion Paper 2021:1.

endowment approach in our case study region. Total work requirements are found to be on average 10% higher in the 'aggregate' approach compared to the 'single plot' approach, while energy requirements are relatively similar (average difference of 2.2%). Among other factors, we find the difference to be highly correlated with the number of plots a farm is endowed with. For instance, the average difference in crop choices increases from the sample average of 11% to 20.8% for those farms that are endowed with less than 10 plots (~ 50% of the case study population). Differences in simulated farm profits when comparing the 'aggregate' land endowment approach to the 'single plot' approach are found to range between -306 €/ha to 434 €/ha (mean: 4.57 €/ha, median: - 9.93 €/ha, S.D.: 71.47 €/ha). Our results suggest that for bio-economic farm analyses focusing on aggregate results over a larger sample of farms, both the 'aggregate' and 'categorized' land endowment approaches are sufficiently accurate in case of similar average numbers of plots per farm as in our study. If single farm results or variability in the population are targeted, we propose to incorporate the 'single plot' approach in bio-economic farm analyses. The same holds for decision support systems focusing on individual farm responses to policy changes or technology adoption.

Keywords

Land aggregation, Land fragmentation, spatial resolution, farm model, BEFM

5.1 Introduction

A farm's land endowment is generally composed of multiple individual plots (Di Falco et al., 2010), defined as the smallest homogeneously managed areas of land in the sense that on each plot one single crop is cultivated (Nesme et al., 2010). Depending on their spatial dispersion, plots may differ in size, soil type and quality, as well as farm-to-field distance. The dispersion of plots over a given area is commonly referred to as land fragmentation (King and Burton, 1982). Higher degrees of land fragmentation are frequent among farming systems around the world, exhibiting both negative and positive consequences in different dimensions (Di Falco et al., 2010;

Geppert et al., 2020; Latruffe and Piet, 2014). While higher land fragmentation fosters biodiversity through crop diversification and increased amount of field margins and hedges (Di Falco et al., 2010; Geppert et al., 2020; Latruffe and Piet, 2014), farm profitability is reduced, as labor requirements and variable costs of cultivation are generally found to be increasing (Di Falco et al., 2010; Janus and Markuszewska, 2017; Latruffe and Piet, 2014; Lu et al., 2018).

Despite these implications, bio-economic farm models (BEFM) rarely consider single plots and resulting indivisibilities in crop choices. Instead, they typically simulate shares of each crop or crop rotation on land endowments, depicted by (in)equality constraints. This neglects possible effects of land fragmentation and does not represent the decision problem faced by farmers, as illustrated by the following example. Suppose a farm is endowed with 15 ha of land divided into three plots of 7.5, 5 and 2.5 ha, on which three possible crops can be cultivated (wheat, rapeseed, and barley). A BEFM depicting the farm's land endowment by a single constraint, and considering additionally maximal crop shares or labor use, may yield optimal crop acreages such as 3.75 ha of rapeseed (25%), 5.625 ha of wheat (37.5%), and 5.625 ha of barley (37.5%). These crop shares cannot be realized without dividing the given plots into smaller units, which may not be feasible or sensible due to technical or management constraints.

Until recently, data on single plots, such as size, soil quality and crop choice, were rarely available as public datasets. BEFMs were therefore forced to model crop choices by shares on aggregate land constraints. However, detailed and spatial explicit plot data become increasingly available for research, for instance based on the digital applications for direct payments under the Common Agricultural Policy which became mandatory in 2016 (European Commission, 2014). To receive financial support, farmers annually report their planned crop choices for each plot based on geo-referenced land registers (cadasters). Such geo-referenced data at plot level can be linked to high resolution maps, for instance on soil type, soil quality or climate (Martini, 2018; Martini et al., 2014). Such increasingly available data allow depicting single plots and related decision taking in BEFMs. The availability of detailed data also increases the potential to use BEFMs in the context of decision support systems (DSS) which aim at supporting farm management decisions. Using a farm's single plots instead of its total land endowment represents more accurately the actual decision problems farmers face (Pahmeyer et al., 2021a) and, thus, potentially increase the acceptance of DSS.

Depicting crop choices on the single-plot level in a BEFM also allows for a better representation of plot related policy measures. Command-and-control instruments as part of agri-environmental policies increasingly prescribe management restrictions depending on a plot's location and further characteristics. For instance, the German implementation of the EU Nitrates Directive comprises restrictions in nitrate sensitive areas at single-plot level. Equally, farmers might specifically enrol plots with lower productivity in agri-environmental opt-in measures. Productivity differences across plots and their consequences for crop choices are also discussed in the literature relating to BEFMs (linear programs). For instance, in his seminal on Positive Mathematical Programming Approach (PMP), Howitt (1995) mentions land heterogeneity as a key reason why linear models with an aggregate land constraint cannot be properly calibrated to observed crop allocation choices.

The simplified modeling of the crop choice problem based on shares of land (type) constraints likely introduces an aggregation bias. The bias is related to plot heterogeneity, i.e. the difference between mean values of plot characteristics as depicted by an aggregate constraint and the values of the individual plots represented by the aggregate. Furthermore, modeling of crop shares on aggregates of land neglects the indivisibility of plots. The magnitude and implications of these two effects have not been studied yet, as it requires a model depicting individual plots and a matching dataset as a benchmark. This paper aims to fill this gap. First, we present and discuss the current state-of-the-art approaches to model crop choices in BEFMs as used for policy and technology evaluation studies and in decision support systems (DSS). Second, we demonstrate how these different approaches affect BEFM model results in a case study consisting of arable farms in the German federal state of North Rhine-Westphalia.

5.2 Material and Methods

5.2.1 *Depicting competition for land in BEFMs*

Our analysis focuses on so-called 'mechanistic' BEFMs which, according to Janssen and van Ittersum (2007), build on existing theory and knowledge, as opposed to 'empirical' BEFMs whose functions are estimated from observed data (Austin et al., 1998). Mechanistic BEFMs are mostly optimization models, frequently based on mathematical programming, either (mixed integer) linear programming (MILP, LP) or quadratic (mixed integer) programming (QMIP, QP) (Janssen and van Ittersum, 2007). Three options to depict the crop choice problem are found in BEFMs as presented in Figure 5.1

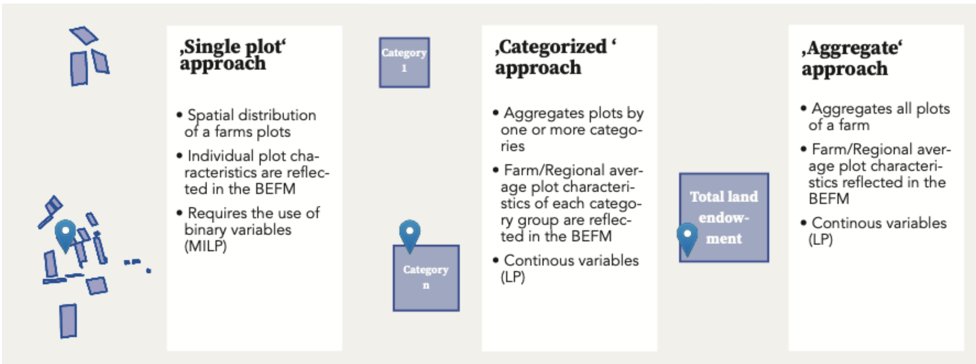


Figure 5.1 Three approaches to depict the land endowment in a BEFM based on mathematical programming.

The first approach depicts the crop choice problem based on a resource constraint relating to a single aggregate land endowment and is therefore referred to as the 'aggregate' (land endowment) approach (Figure 5.1, right panel). Accordingly, the sum of the cultivation areas X_j (in ha) of crops j is required to be less than the total land endowment b. Given gross margins of each crop c_j (in €/ha), a simple, total gross margin (Z) maximizing farm LP may be written as follows (following the notation from Hazel and Norton (1986)):

$$\max Z = \sum_{j}^{n} c_j \cdot X_j \tag{5.1}$$

such that

$$\sum_{j}^{n} X_j \leq b \tag{5.2}$$

and

$$X_j \geq 0, \text{for all } j = 1 \text{ to } n \tag{5.3}$$

In this approach the cultivation area designated to a certain crop (X_j) is given as a share of the aggregate land endowment b.

The second approach, referred to as the 'categorized' (land endowment) approach, extends the first by disaggregating the total land endowment into sub-categories. The land endowment can for instance be differentiated by type of land (arable, grassland), soil type, soil-climate-zone or a combination of these. For each subcategory of land s, different sets of allowed crops j_s may be defined and gross margins for each crop might differ across land sub-categories, i.e. $c_{j,s}$. Incorporating these changes, the LP depicted by Eq. 5.1 – Eq. 5.3 may be extended as follows:

$$\max Z = \sum_{j}^{n} \sum_{s}^{o} c_{j,s} \cdot X_{j,s} \tag{5.4}$$

such that

$$\sum_{j}^{n} X_{j,s} \leq b_s \text{, for all } s = 1 \text{ to } o \tag{5.5}$$

and

$$X_{j,s} \geq 0, \text{for all } j = 1 \text{ to } n \tag{5.6}$$

Both approaches apply the same modeling principle of designating a fraction of (a subcategorized) land endowment to a certain crop, rendering $X_{j,s}$ or X_j positive, continuous variables.

The third approach considers single plots by using binary variables instead. Gross margins for each crop j can now be differentiated for each plot k (figure $c_{j,k}$, in €/ha). A binary variable $V_{j,k}$ indicates whether crop j is selected (=1) or not (=0) for plot k. The gross margin realized on a plot is the plot specific gross margin $c_{j,k}$ per ha of the selected crop times the plot size x_k in ha. The introduction of the binary variables $V_{j,k}$ leads to a so-called 'binary integer programming' or 'mixed-integer programming', the latter if the BEFM also contains continuous variables.

The resulting (mixed) integer program, referred to as the 'single plot' (land endowment) approach, may be written as follows:

$$\max Z = \sum_{j}^{n} \sum_{k}^{m} c_{j,k} \cdot x_k \cdot V_{j,k} \tag{5.7}$$

such that

$$\sum_{j}^{n} V_{j,k} = 1, \text{for all } k = 1 \text{ to } m \tag{5.8}$$

and

$$V_{j,k} \in \{0,1\}, \text{for all } j,k = 1 \text{ to } n,m \tag{5.9}$$

The 'categorized' approach could allow for plot specific analyses if each plot received its own land sub-category s. However, as this approach uses continuous variables, it returns optimal shares of crops on each plot ($X_{j,s}$) and implies that plots may be split arbitrarily. We do not consider this further as splitting plots breaks their definition as the smallest homogeneously managed units of land.

The differences in the simulation results using the 'aggregate' or 'categorized' approach compared to the 'single plot' approach relate to two main effects. First, the aggregation bias resulting from the aggregation over plot characteristics as a measure of land fragmentation (plot size, soil quality, farm-to-field distance). In the case of the 'categorized' approach, the aggregation bias will largely be driven by the number of categories, and whether the model results are sensitive to the choice of categorization (e.g. categorization by soil type, soil quality, single plots). Second, the effect of considering indivisibility in the 'single plot' approach compared to the fractions allowed in the 'aggregate' and 'categorized' approach. Here, the assumption that plots refer to the smallest homogenously managed units of land plays a central role, as this implies that the plots cannot be divided into smaller sub-units in our analysis.

Table 5.1 gives examples of BEFMs identified from the literature for each of the three approaches. As noted by Janssen and van Ittersum (2007), many BEFMs are developed for specific case studies and are rarely reused. For the sake of simplicity and relevancy, the overview presented in Table 5.1 is limited to some frequently used BEFMs in Europe, mainly drawing on the review article of Britz et al. (2021).

Approach	*Used by (selection of BEFMs)*	*Primary use cases*
'Aggregate'	**CAPRI-FT** (Gocht et al., 2017, 2013; Gocht and Britz, 2011; Schroeder et al., 2015)	Regional/Sectoral policy analysis
	IFM-CAP (Louhichi et al., 2018, 2015; M'barek et al., 2017)	
'Categorized'	**FSSIM** (Kanellopoulos et al., 2014; Louhichi et al., 2010; van Ittersum et al., 2008)	Ex-ante on-farm analysis of policy and technology adoption
	ORFEE (Mosnier et al., 2017)	
	FARMDYN (Kuhn et al., 2019, 2020; Lengers, 2012; Lengers et al., 2014, 2013; Pahmeyer and Britz, 2020; Seidel and Britz, 2019)	
'Single plot'	**MINRISK** (Radulescu and Radulescu, 2012)	Decision support
	FRUCHTFOLGE (Pahmeyer et al., 2021a)	

Table 5.1 Use of the different land endowment approaches in the literature, mainly based on Britz et al. (2021).

5.2.2 *Design of experiments*

For the underlying analysis, the BEFM FRUCHTFOLGE (Pahmeyer et al., 2021a) is used to examine how the different land endowment approaches affect the simulation results. FRUCHTFOLGE is chosen as it incorporates the technically demanding 'single plot' approach as its default. As the 'categorized' and 'aggregate' approaches are simplifications of the 'single plot' approach, these can be modeled in the FRUCHTFOLGE BEFM without requiring changes to the codebase of the model. Technically, we redefine the $V_{j,k}$ as fractional variables and define one or multiple categorized larger plots, which depict the average characteristics of the single plots and their summed-up size. FRUCHTFOLGE is an open-source software, and available in a public code versioning repository[2].

The 'categorized' land endowment approach allows for varying level of detail. Considering the BEFMs outlined in Table 1, all models distinguish between arable and permanent grassland, and some additionally between soil types (ORFEE and FARMDYN). According to the focus of this paper, only arable farms without livestock are considered to isolate the effects of the varying plot characteristics and land endowment approaches on the results. Therefore, differentiation between arable and permanent grassland is not used in the 'categorized' approach, instead we depict the more evolved differentiation by soil type.

The arable farms are given the option to cultivate nine of the most frequently cultivated crops in the case study area, jointly accounting for more than 78% of the total arable land (IT.NRW, 2019). Prices and direct costs (seeds, fertilizers, plant protectants) for each crop represent averages of the past 18 years within the case-study region (KTBL, 2020). Plot specific yields are calculated based on a linear regression function including the soil quality as an independent variable[3], estimated from average yields and soil quality ratings in the 45 NUTS 2 regions in Germany over 19 years. In the

---

[2] The code is hosted at the following GitHub repository: https://doi.org/10.5281/zenodo.4765941

[3] See the following notebook for details: https://observablehq.com/@chrispahm/influence-of-soil-quality-and-soil-moisture-index-on-crop-yi. The regression results can also be found in the appendix, Table A3.

‘categorized’ approach, crop yields are calculated for each soil type category of the farm, using the average soil quality of the plots within the category. In the ‘aggregate’ approach, whole farm average yields are calculated based on the average soil quality of the plots.

Machine costs are calculated using the regression model from Heinrichs et al. (2021) which considers farm-to-field distances and plot sizes. In the ‘single plot’ approach, the individual farm-to-field distance and plot size of a plot are reflected in the calculations. In the ‘categorized’ approach, average farm-to-field distances and plot sizes for each soil type category are considered, while in the ‘aggregate’ approach, whole farm averages are taken. For all crops, a fixed gross wage rate of 19.19 €/h (net wage rate of 13.5 €/h) is assumed (KTBL, 2020). The calculation of the labor costs per crop follows the same concept as the calculation of the machine costs for each land endowment approach.

The profitability per ha for a crop is calculated as the difference between crop revenues and direct costs as well as costs for machinery and labor. Figure 5.2 illustrates the resulting difference between the ‘aggregate’ and the ‘single plot’ approach using the example of winter wheat. In the ‘aggregate’ and ‘categorized’ approach, the profitability of a crop is independent of the chosen share, reflecting the constant returns to scale of the technology underlying the Leontief production function used in a LP. This is not the case for the ‘single plot’ approach. Here, the average realized profit per ha of a crop changes depending on which plot the crop is cultivated on. Ordering the plots from highest to lowest profitability in Figure 2 shows that this implies decreasing return to scale, similar to the convexity found in quadratic programming approaches typically used with PMP (Heckelei et al., 2012).

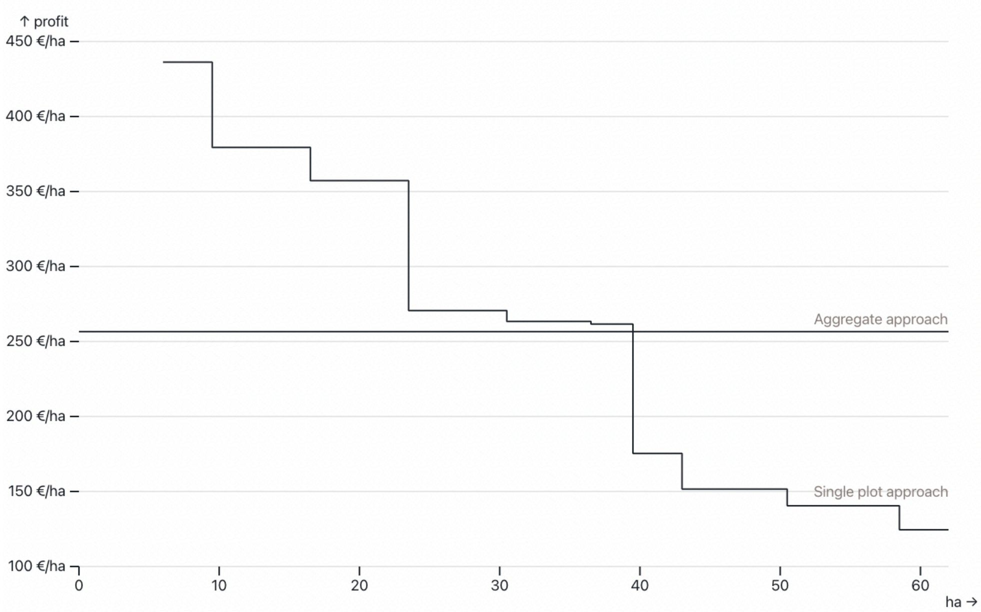


Figure 5.2 Marginal profitability per hectare of wheat cultivation for all plots of an exemplary farm in both the 'aggregate' (straight line), and 'single plot' (stepped line) approach, sorted by descending order. Plots exhibiting the highest marginal profitability are generally characterized by higher soil qualities, larger plot sizes, and closer proximity to the farm.

Table 5.2 gives an overview of the input-output coefficients for each crop, including minimum and maximum values for the case study region, reflecting varying soil qualities, farm-field-distances and plot sizes[4].

[4] The data may be explored interactively in the following notebook: https://observablehq.com/@chrispahm/crop-gross-margins-in-germany?collection=@chrispahm/agriculture/2

Crop	*Price (€/dt)*	*Yield (dt/ha)*	*Revenues (€/ha)*	*Costs (€/ha)*	*Profit (€/ha)*
Field beans	18.52	26.57 45.02	492.08 833.77	820.5 1315.3	-823.2 13.3
Wheat	16.95	58.61 97.35	993.44 1650.08	986.6 1656.2	-662.8 663.5
Rye	15.69	47.85 66.38	750.77 1041.50	899.1 1410.7	-659.9 142.4
Barley	16.01	44.87 85.48	718.37 1368.53	925.4 1508.2	-789.9 443.2
Maize - Corn	17.38	79.36 113.38	1379.28 1970.54	1562.6 2109.4	-730.1 407.9
Rapeseed	36.53	30.99 44.82	1132.06 1637.27	1020.9 1529.6	-397.6 616.4
Sugar beets	3.54	675.68 814.10	2391.91 2881.91	1404.1 1879.7	512.2 1477.8
Maize - Silage	2.80	388.18 510.22	1086.90 1428.62	1315.4 2093.2	-1006.3 113.2
Summer oats	15.13	36.75 58.87	556.03 890.70	709.0 1209.0	-652.9 181.7

Table 5.2 Economic figures for each crop allowed to be cultivated in the BEFM. If present, multiple rows per column indicate minimum (top row) and maximum (bottom row) values. Data based on KTBL (2020) and Heinrichs et al. (2021).

Note: For the underlying minimum and maximum values of soil quality, plot size and farm-to-field distance, see the following section (section 2.3).

Constraints controlling maximum allowed crop shares are introduced in the BEFM for all three land endowment approaches, they reflect minimum waiting period between years where the same crop is cultivated on a plot (see Table A1 in the appendix). Due to the agronomic intolerance of sugar beets and rapeseed in crop rotations, their combined maximum share is limited to 33% (ISIP, 2021). Further constraints reflect obligations from the EU's so-called "greening" measures: Farms above 10 ha and below 30 ha need to cultivate at least two crops, with the major crop not covering more than 75% of the arable land. Farms above 30 ha need to cultivate a minimum of three crops, with the major crop not covering more than 75%, and the sum of the two major crops not covering more than 95% of the arable land. Furthermore, farms endowed with more than 15 ha need to devote 5% of their arable land to a so-called ecological focus area. For the farms affected by this measure, the constraint needs to be fulfilled by cultivating 5% of field beans in our simplified model. A detailed description of the greening measures is provided by Gocht et al. (2017).

Each farm is simulated once for each of the three land endowment approaches and subsequently, results of the 'categorized' and 'aggregate' approaches are compared with the results of the 'single plot' approach. The provided indicators depict agronomic ('Summed difference in crops shares'), social ('Difference in total work load'), environmental ('Difference in cumulative energy requirement'), and economic ('Difference in profit per ha') differences. The 'Summed difference in crops shares' indicator for a farm is calculated as follows. First, the absolute differences of the area allocated to each crop j under the aggregate approaches (X_agg_j, both for the 'categorized' and 'aggregate' approaches) and the 'single plot' X_bin_j approach are summed up. Second, to account for farm size, the resulting sum is divided by the farm's total land endowment b. And third, as a deviation in the share for a crop implies a deviation in the opposite direction for other crops, the average absolute deviation it divided by two:

$$\text{Summed difference in crop shares (\%)} = \frac{1}{2}\sum_{j}^{n} \frac{\left|X\_bin_j - X\_agg_j\right|}{b} \tag{5.10}$$

Following this calculation, the 'Summed difference in crop shares' indicator results in a percentage value defined in the range $[0{,}100\%]$.

In order to evaluate which farm characteristics drive differences of indicator results, ordinary least squares (OLS) regressions are performed in R, Version 3.6.1 (R Core Team, 2019).

5.2.3 *Case study region*

Our analysis builds on a synthetic farm population of the German federal state of North-Rhine Westphalia (Pahmeyer et al., 2021b) of which all 8,509 specialized arable farms are considered. As the classification is based on shares of revenues by farm branch in total farm revenues (Kuhn and Schäfer, 2018), specialized arable farms might still be involved, for instance, in fattening of ruminants and manage some grasslands. However, the management of permanent grassland is left out of the analysis as animal husbandry is not considered. Figure 5.3 gives an overview of the spatial dispersion of the main farm characteristics such as average soil quality depicted by the "Muencheberg soil quality rating" (SQR) (Mueller et al., 2014), farm-to-field distances, plot sizes and number of plots per farm in North-Rhine Westphalia. Figure 5.4 gives an overview of the distribution of these farm characteristics among the population. The arable land endowments of the farms range between < 1 ha to 490 ha. The mean farm size is 42 ha, the median farm size is 23 ha. The farms' average SQRs range from 23 to 95, with a median value of 68 (mean: 68). The deviation of SQR values within a farm is up to 42.40, with a median value of 4.04 and a mean value of 7.05. Average field-to-farm distances range from 0.07 km to 17.77 km (median: 0.84 km, mean: 1.33 km). The average standard deviation (S.D.) of the farm-to-field distances is 1.29 km (median: 0.59 km). Plot sizes range from 0.06 ha to 33.20 ha (median: 2.5 ha, mean: 2.76 ha), and have an average S.D of 2.24 ha (median: 1.94 ha). The number of plots the farms are endowed with range between 1 and 202, with a mean of 14 plots and a median value of 9 plots.

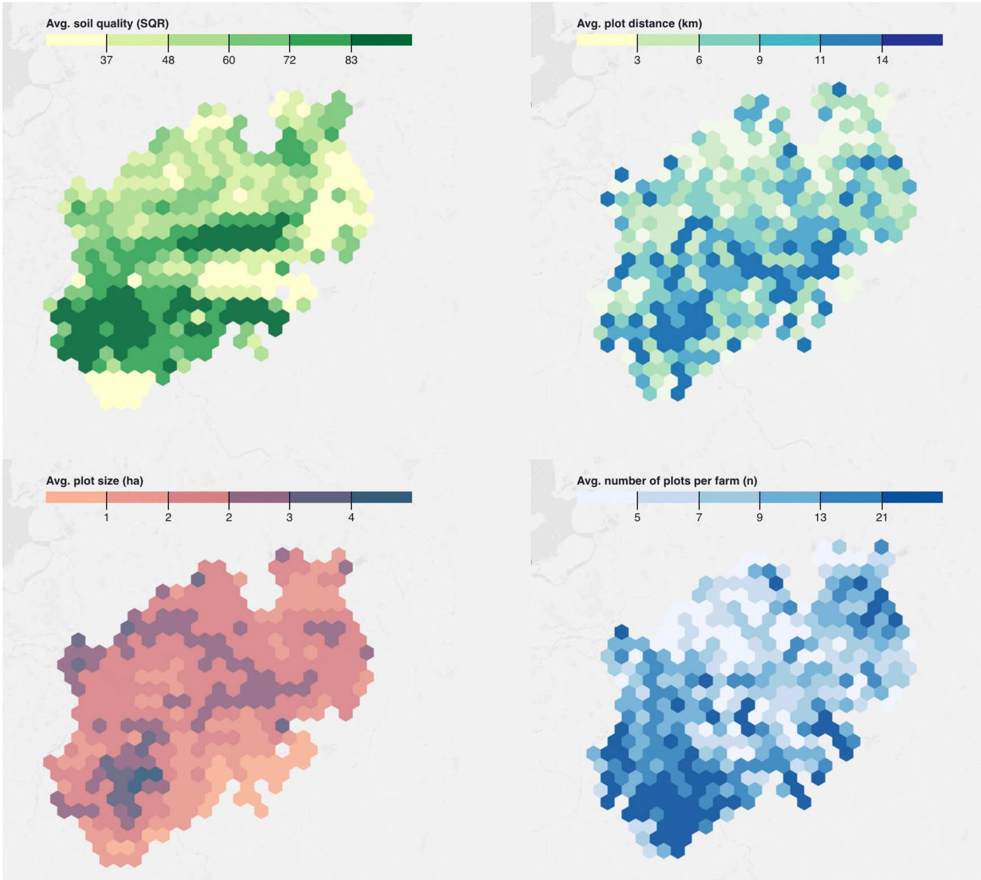


Figure 5.3 Distribution of average soil quality ratings (SQR), plot (farm-to-field) distances, plot sizes, and number of plots per farm within the case study region of North Rhine-Westphalia.

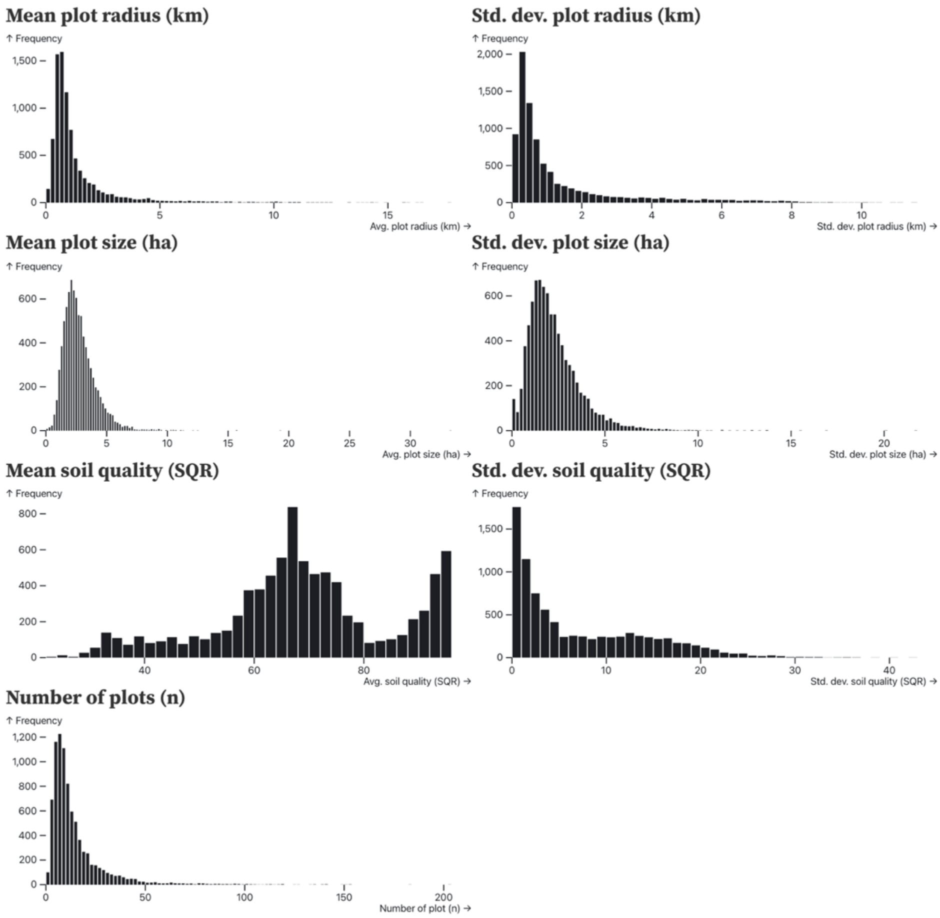


Figure 5.4 Histograms of selected farm characteristic among the case-study farm population.

5.3 Results

Simulation results for the four considered indicators under the ‘aggregate’ and ‘categorized’ approach are found to be largely equal. The crop allocation results between the two approaches are the same for 99.8% of the simulated farm population. Also, the farm profits (€/ha) are found to be largely equal for a greater part of the population, with 75% of the farms expressing a difference of less than 9 €/ha and 95% of the farms expressing a difference of less than 39 €/ha. A detailed comparison of the indicator results of the ‘aggregate’ and ‘categorized’ approach is shown in Figure A1 in the appendix.

For brevity, the results section therefore focuses on the difference between the ‘aggregate’ and ‘single plot’ approach, as depicted in Figure 5.5. On average, the summed difference in crop shares between the two approaches is 11.15% (median: 2.23%, S.D.: 19.65%). The average difference in workload is 10.8%, while the median is 7.56% (S.D.: 11.86%), i.e., the ‘aggregate’ approach overestimates the required labor needs in the sample. The opposite is found for the cumulative energy requirement which is on average 2.23% lower in the ‘aggregate’ approach compared to the ‘single plot’ approach (median: 0.4%, S.D.: 7.35%). The simulated average farm profits are found to be slightly higher in the ‘aggregate’ approach when compared to the ‘single plot’ approach (4.57 €/ha, median: -9.93 €/ha, S.D.: 71.47€ /ha). This effect is likely related to the relaxation of the indivisibility underlying the ‘aggregate’ approach, and the corresponding different crop shares. Despite the overestimated labor needs and thus costs in the ‘aggregate’ approach, which go along with higher machinery hours and costs, the average farm profits are on a par with the ‘single plot’ approach. This implies that the share of crops with larger revenues is higher under the aggregate approach which can also be seen in Figure 5.7. The histograms in Figure 5.5 reveal that the differences of the indicator values are found to be relatively small for a high share of farms. However, for a small part of the population, the simulation approaches show large differences for the indicators, especially for farm profits.

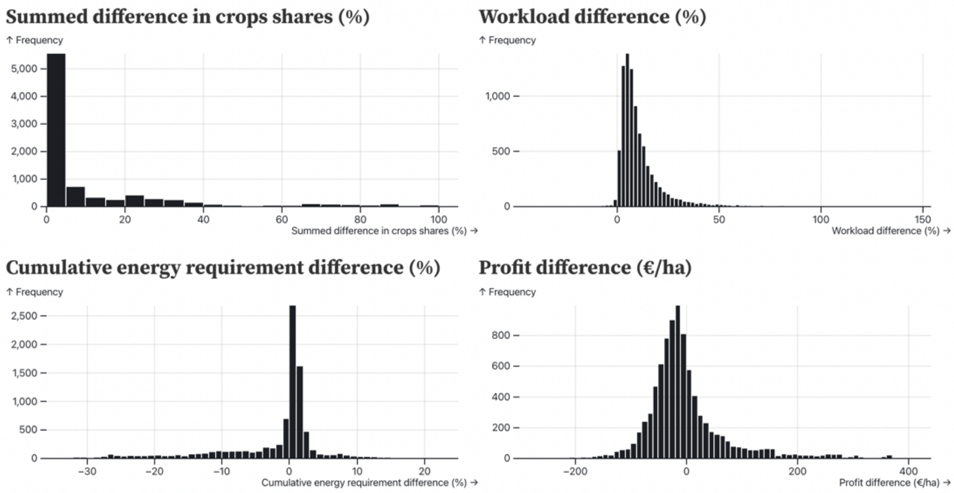


Figure 5.5 Histograms of differences in indicator levels comparing the simulations results of the 'aggregate' approach with the results of the 'single plot' approach in the farm population.

The indicator values resulting from the comparison of the ‘categorized’ approach with the ‘single plot’ approach are very similar to the results of the ‘aggregate’ approach, see Figure A1 in the appendix.

In order to illustrate how the different land endowment approaches combined with the varying characteristics of land fragmentation lead to differences in the crop shares, Figure 5.6 displays the crop choices resulting from the different land endowment approaches for an exemplary farm endowed of 12.07 ha. Since all of the farm’s plots are of the same soil type, the simulation results of the ‘aggregate’ and ‘categorized’ approach are the same for this farm. However, note that plots are still heterogeneous considering their soil quality and field-to-farm distance. While the crop shares of wheat and rapeseed are also mainly similar between the ‘single plot’ and the ‘aggregate’ approach, larger differences are found for rye and maize. Considering the farm’s average soil quality, field-farm-distance, and plot size, the average profit of cultivating maize is -223.66 €/ha, while it is -309.95 €/ha for rye. In order to maximize profits (or minimize losses in this case), maize is selected over rye in the ‘aggregate’ (and ‘categorized’) approach. Despite their same soil type, the plots in the far east of the farm exhibit a very low soil quality of 26 (SQR) (farm median: SQR of 69). On these plots, the losses of -436.27 €/ha for rye are smaller than for maize with -476.95 €/ha, reversing the order compared to the average. The BEFM therefore selects rye on these fields instead of maize in the ‘single plot’ approach. Given that the farm is endowed with 20 plots, the indivisibility of plots can largely be disregarded as a factor influencing the crop share differences, as many different combinations of plots are present to come close to a desired crop share. The example rather shows how the aggregation bias from the ‘aggregate’ approach is caused by differences between average and plot specific values of plot characteristics.

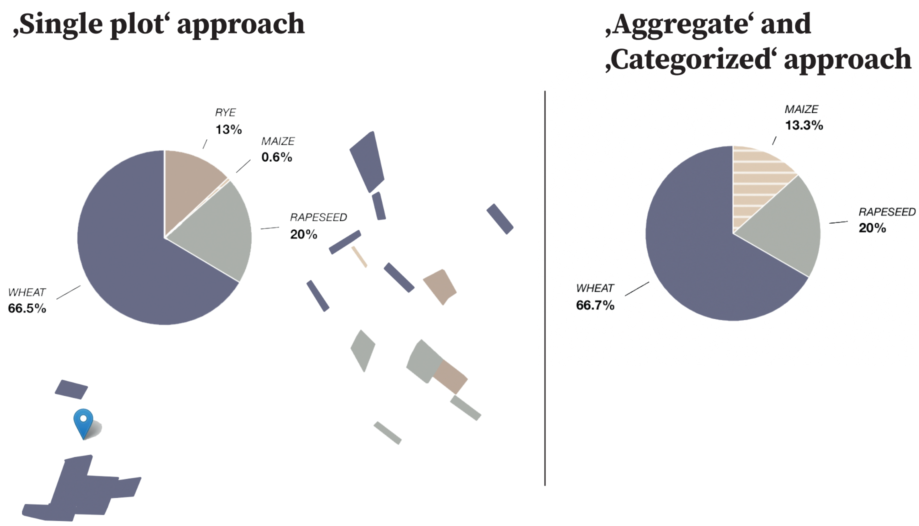


Figure 5.6 Crop choice results of the three land endowment approaches for an exemplary farm. The blue marker in the left panel displays the farms location.

Table 5.3 presents the results of standardized multiple linear regression models (OLS) on differences of the four chosen indicators between the ‘aggregate’ and ‘single plot’ approach. The similar results for the ‘categorized’ approach can be found in the appendix (Table A2). For (highly) auto-correlated land fragmentation and farm characteristics (e.g., farm size and number of plots, or mean and median values of the same parameter), the characteristics with the highest-ranking Pearson’s correlation coefficient are used in the regression models (Figure A2 in the appendix). The number of plots is log transformed due to the stronger influence of fewer plots on the indicator values.

Table 5.3 suggests that the number of plots present on a farm is the main driver for the 'Summed difference in crop shares' indicator. Farms endowed with fewer plots generally express greater differences in the optimal crop allocation between the two approaches, which reflects the impact of the assumed indivisibility of the plots. Furthermore, also the S.D. in the plot sizes of the farm, the mean plot size, as well as plot radii are found to have a stronger influence on the difference of the results in this indicator, displaying the influence of these factors on the aggregation bias. While higher values of the S.D. in plot sizes, as well as higher mean plot radii are found to increase the overall difference in crop allocation results, higher mean plot sizes, as well as S.D. in plot radii decrease the difference.

Considering the difference in workload among the different land endowment approaches, again the number of plots, but also the mean- and S.D. of plot sizes within a farm are found to have a stronger influence. While the S.D. of plot sizes is found to increase the difference in the workload simulation results, both an increasing number of plots as well as an increasing mean plot size are found to decrease the difference in the simulation results.

Also, the difference in profit per ha between the two land endowment approaches is mainly influenced by the number of plots (effect of indivisibility), followed by the mean- and S.D. of plot sizes, and the farms S.D. in farm-to-field distances (aggregation bias). While larger mean plot radii and plot sizes per farm tend to have a positive influence on the profit difference (higher profits in the 'aggregate' approach simulation results compared to the 'single plot' approach), the number of plots, S.D. in plot radii, as well as the S.D. in plot sizes have a negative influence on the profit difference (higher profits in the 'single plot' approach, compared to the 'aggregate' approach).

Similar to the differences in profit, also the differences in the cumulative energy requirement (CER) are mostly depending on the number of plots, as well as the farms mean- and S.D. of plot sizes. In this indicator, the number of plots, mean plot size, as well as the S.D. in plot radii is found to have positive influence on the difference in simulation results. On the other hand, the farms mean plot radius, S.D. in plot sizes, as well as the mean soil quality is found to have a negative impact on the difference in CER simulation results.

	Dependent variable:			
	OLS			
	'Aggregate' vs 'single plot' approach			
	Summed diff. in crop shares (%)	Diff. workload (%)	Diff. profit (EUR/ha)	Diff. CER (%)
Mean plot radius farm [km]	0.150***	-0.049***	0.122***	-0.158***
Dev. plot radius farm [km]	-0.094***	0.111***	-0.130***	0.108***
ln(Number of plots [n])	-0.674***	-0.232***	-0.584***	0.575***
Mean plot size farm [ha]	-0.156***	-0.475***	0.205***	0.146***
Dev. plot size farm [ha]	0.303***	1.209***	-0.304***	-0.188***
Mean soil quality farm [SQR]	0.027***	0.021***	0.060***	-0.160***
Dev. soil quality farm [SQR]	0.040***	0.008	-0.007	0.057***
Constant	-0.035***	-0.022***	-0.010	0.027***

Observations	8,409	8,409	8,409	8,409
R^2	0.474	0.741	0.420	0.350
Adjusted R^2	0.474	0.740	0.420	0.349
Residual Std. Error (df = 8401)	0.651	0.510	0.716	0.748
F Statistic (df = 7; 8401)	1,081.725***	3,424.726***	870.023***	645.777***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.3 Standardized regression results (beta coefficients) for different indicators comparing the BEFM simulations results from the 'aggregate' and 'single plot' land endowment approach.

Figure 5.7 displays the summed cultivation area in the farm population resulting from the simulation of the ‘aggregate’ and ‘single plot’ approaches. Mainly due to the indivisibility effect, the total cultivation area of the more profitable crops, namely wheat, sugar beets, as well as winter rape is higher in the simulation results of the ‘aggregate’ approach compared to the ‘single plot’ approach. On the other hand, the cultivation area of field beans, (corn) maize, oats, silage maize, winter barley, and winter rye is higher in the ‘single plot’ simulation results.

In the ‘aggregate’ approach, 64% of the total farm population hit the maximum crop share constraint for sugar beets (20% max. crop share), and 36% of the population do so for winter rape (also 20% max. crop share). Due to the indivisibility of plots, these figures cannot be reliably calculated for the ‘single plot’ approach.

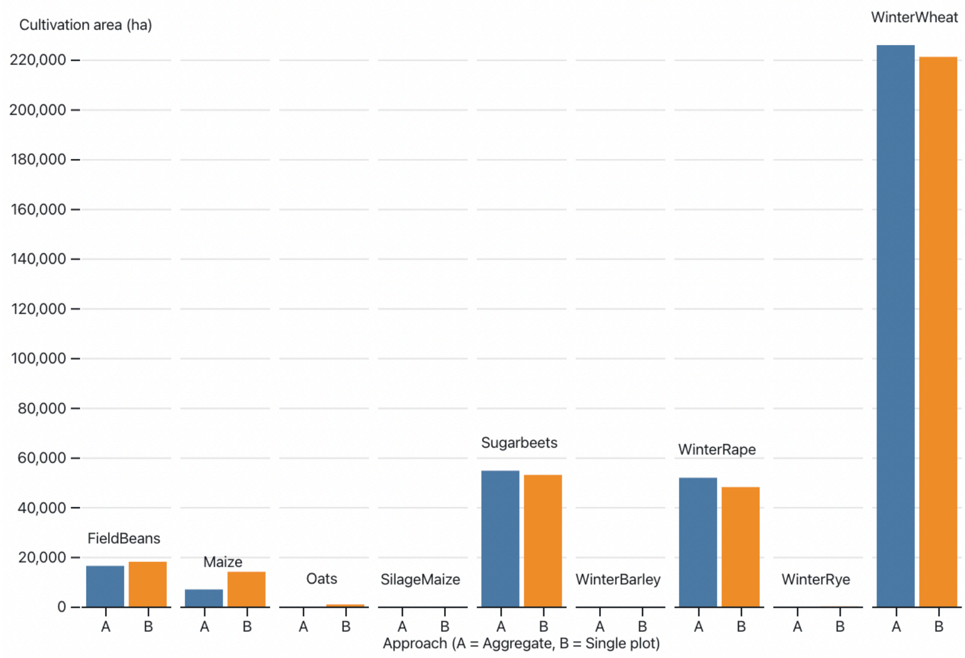


Figure 5.7 Total simulated cultivation area per crop in the aggregate (A, blue) and single plot (B, orange) approach.

Table 5.4 displays the standardized regression coefficients measuring the impact of various farm characteristics on the relative difference in the total cultivation area for each crop between the ‘aggregate’ and ‘single plot’ approach. For all crops with a notable difference in the summed total cultivation area between the two approaches (see Figure 7), the number of plots as a measure of the indivisibility effect, the mean plot size, as well as the intra-farm S.D. of plot sizes are found to have the greatest influence on the relative difference in the crop shares. For oats, also the mean, as well as the S.D. of soil qualities is found to have an impact on the relative difference in the total cultivation area.

	Dependent variable:
	Aggregate vs 'single plot' (rel. diff)
	OLS

	Field beans	Wheat	Rye	Barley	Maize - Corn	Rape	Sugar beets	Maize - Silage	Oats
Mean plot radius farm [km]	-0.059***	-0.138***	0.001	-0.006	0.154***	-0.130***	-0.087***	0.007	0.012
Dev. plot radius farm [km]	0.016	0.070***	-0.028	-0.008	-0.085***	0.106***	0.078***	-0.010	-0.023
ln(Number of plots [n])	-0.169***	0.555***	-0.069***	-0.056***	-0.621***	0.657***	0.459***	-0.024***	-0.113***
Mean plot size farm [ha]	0.439***	0.311***	0.108***	0.050**	-0.209***	-0.157***	-0.187***	-0.035**	0.009
Dev. plot size farm [ha]	-0.195***	-0.434***	-0.070***	0.144***	0.331***	0.037***	0.072***	0.027**	-0.049***
Mean soil quality farm [SQR]	-0.011	-0.001	-0.099***	0.016	0.092***	-0.061***	-0.109***	-0.030***	-0.185***
Dev. soil quality	0.035***	-0.005	0.085***	0.001	-0.015*	-0.028***	-0.069***	0.004	0.120***

farm [SQR]									
Constant	0.018*	0.041***	0.008	-0.0001	-0.037***	0.013	0.00001	-0.016**	0.008
Observations	8,409	8,409	8,409	8,409	8,409	8,409	8,409	8,409	8,409
R^2	0.095	0.356	0.029	0.033	0.415	0.405	0.198	0.005	0.079
Adjusted R^2	0.094	0.355	0.028	0.032	0.415	0.405	0.197	0.004	0.078
Residual Std. Error (df = 8401)	0.954	0.724	0.982	0.990	0.685	0.732	0.855	0.676	0.962
F Statistic (df = 7; 8401)	125.849***	662.730***	36.147***	40.423***	852.565***	818.594***	296.283***	5.533***	102.358***
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$								

Table 5.4 Standardized regression results (beta coefficients) comparing the relative difference in crop cultivation area for each crop between the BEFM simulations results from the 'aggregate' and 'single plot' land endowment approach.

5.4 **Discussion**

Our results empirically quantify the effects of different aspects of land fragmentation on the simulation results of a mechanistic BEFM. The regression analysis shows that the smaller the number of plots, the larger the differences between the binary choice model with plots depicting land heterogeneity and the LP which optimizes crop shares under a constraint assuming homogenous land. This suggests that the effect of indivisibility dominates over the aggregation bias. As the aggregation bias rather increases with growing numbers of plots in a farm, the opposite effect would be found in the regression analysis if the aggregation bias was the major driver of differences.

Most of the indicator values tested in our study are centered around a mean difference being close to zero (see Figure 5). Therefore, analysis focusing on findings for a whole farm population will likely attain similar average results between the 'single plot' and the 'aggregate' land endowment approach. However, as seen from the relatively wide range in indicator values, especially for the profitability of crop cultivation per hectare, simulation results for the selected farms under the two approaches can differ substantially. Therefore, for studies focusing on selected case study farms and their responses to new policies or technologies, either the 'single plot' approach or the 'categorized' approach using a sufficient number of categories is recommended. In the context of DSS however, solely the 'single plot' approach is recommended as it depicts the decision problem farmers face more accurately (see Pahmeyer et al., 2021a). Furthermore, as the 'single plot' approach also considers the actual required workload for each specific plot, compared with farm averages over all plots, this approach is deemed more appropriate in a decision support context. However, it has to be considered that such heterogeneity requires integer crop choices, which renders model calibration far more difficult (Britz, 2021) compared to established approaches such as PMP (Heckelei et al., 2012). Equally, using integers to depict crop choices increases the 'jumpiness' in the allocative responses, and the overall higher model detail also implies that model result interpretation is rendered more demanding.

In our dataset, the shares of the three dominant crops (winter wheat, sugar beet and rape seed) are mostly driven by maximal crop rotational constraints, which means that their profitability advantages over other crops do not (much) depend on soil quality, plot size or farm-to-field distance. Output price fluctuations for crops are not necessarily highly correlated, take sugar beets and cereals as an example. Hence, crops might be found as dominant or not depending on the considered years when calculating the profitability of each crop. The importance of this effect is therefore likely case-study dependent. The closer the profitability of crops are to each other under average plot characteristics of a farm, and the larger the heterogeneity of the plots, the more likely it is to find aggregation bias in the 'aggregate' and 'categorized' approach in relation to the optimal crop shares.

In order to reproduce empirically observed crop shares in the baseline model results, BEFMs are commonly calibrated using either PMP (Heckelei et al., 2012) or by some more or less automated approach to adjust coefficients in MILPs or LPs (Britz, 2021). Among others, Howitt (1995) states heterogeneous land quality and the corresponding variations in crop yields as a likely reason for the need of calibration, such that linear models are not well suited to recover observed crop allocation changes.

No attempt is made here to calibrate the three competing modelling approaches which, if successful, would remove the differences at least with regard to crop choices. Our findings certainly do not imply that an integer-based, normative crop choice model depicting single plots generally leads to allocative responses more closely resembling empirically observed crop shares. It is however clear that its calibration against observed allocative responses is more demanding (Britz, 2021), whereas PMP based models using crop shares can be calibrated relatively straightforward against given price elasticities (Mérel and Bucaram, 2010). We also assume in all three models that labor is bought (or sold) at a given price, by considering its costs in the profitability per hectare. A BEFM might instead comprise annual or sub-annual labor constraints, which likely restrict the solution space further and thus potentially reduce differences between the three modelling approaches. However, these points mostly apply to BEFMs being used in a positive, policy or technology evaluating context. BEFMs used for decision support are generally not calibrated to empirically observed crop shares, as they aim to explore optimal solutions to the allocation problem given a farm specific, constrained set of resources, and therefore do not aim to predict farmers behavior (Reidsma et al., 2018).

In our analysis, the profitability of a crop solely depends on plot attributes, not on farm or farmer's characteristics. This allows analyzing impacts of the aggregation bias caused by plot heterogeneity and indivisibility independently of other effects. In empirical analysis, especially farm size is likely closely correlated with the number of plots present in a farm. This makes it harder to disentangle effects of the number of plots and plot heterogeneity from effects of farm size. Farm size likely affects crop profitability and crop choice, for instance, by size depending on differences in the costs of depreciation, in transaction costs, or in mechanization level. Such effects are not considered in our analysis. Farm size also likely affects farmer's behavior, such as via impacts of wealth on risk behavior (Sulewski et al., 2020), whereas our models assume risk neutrality. More generally, the importance of plot indivisibility for crop choices challenges the usual assumption on differentiable functions and error term distributions in empirical work in this field.

Note that differences between the simulation results of the different land endowment approaches reported in this manuscript assume an ideal parameterization for each farm. Lacking farm specific information, many BEFMs only adjust prices and sometimes yield levels for individual farms, and use regional averages for other parameters, such as variable costs of crop production. For instance, recent studies applying BEFMs to German farms use a farm-field-distance of 2 km and a plot size of 2 ha defined as the default values found in planning data collections (Kuhn et al., 2020; Lengers et al., 2014; Pahmeyer and Britz, 2020; Schäfer et al., 2017). The difference in the simulation results between the 'aggregate' land endowment approach using such default values compared to the results of the 'single plot' is higher than reported in our manuscript, as our analysis still reflects in the aggregate approaches farm specific plot averages.

5.5 Summary and Conclusion

The aims of our manuscript are to identify approaches to model crop choices in BEFMs and to quantify differences in their results, based on a case study consisting of arable farms in the German federal state of North Rhine-Westphalia. Our findings suggest that results may vary substantially between the approaches. While we find quite limited differences between the 'aggregate' and 'categorized' approaches, their results are systematically different from the 'single plot' approach. The results of a regression analysis suggest that differences are mainly driven by the number of plots a farm is endowed with, while other characteristics such as the intra-farm S.D. of soil qualities, plot sizes, and driving distances show a significant, but less relevant influence. Thereby, the indivisibility of plots is the major driver for the differences in our results. Accordingly, the heterogeneity of plots and the corresponding aggregation bias is of minor importance in our analysis.

Following our simulation results, we suggest that both the 'aggregate' and 'categorized' land endowment approaches yield sufficiently accurate results for studies involving policy analysis or technology adoption for a whole farm population. For BEFMs used in policy and technology analysis, effects of plot heterogeneity can likely be considered by a sufficiently large number of land categories in the 'categorized' approach, and our analysis suggests that especially soil quality differences can be relevant here.

Recommendations are likely different for BEFMs targeting single farm results or variability in the farm population, as well as DSS. Considering the wide range of profit differences between the 'single plot' and 'aggregate' approach among the population, studies targeting single farms and DSS should incorporate spatially explicit, single plots to better capture the decision problem and provide accurate decision support for every individual farm.

5.6 Acknowledgements

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2070 – 390732324.

5.7 References

Austin, E.J., Willock, J., Deary, I.J., Gibson, G.J., Dent, J.B., Edwards-Jones, G., Morgan, O., Grieve, R., Sutherland, A., 1998. Empirical models of farmer behaviour using psychological, social and economic variables. Part I: linear modelling. Agricultural Systems 58, 203–224. https://doi.org/10.1016/S0308-521X(98)00066-3

Baeumer, K., 1990. Gestaltung der Fruchtfolge, in: Integrierter Landbau. BLV Verlagsgesellschaft mbh, Munich.

Britz, W., 2021. Automated Calibration of Farm-Sale Mixed Linear Programming Models using Bi-Level Programming (in press). German Journal of Agricultural Economics 70.

Britz, W., Ciaian, P., Gocht, A., Kanellopoulos, A., Kremmydas, D., Müller, M., Petsakos, A., Reidsma, P., 2021. A design for a generic and modular bio-economic farm model. Agricultural Systems 191, 103133. https://doi.org/10.1016/j.agsy.2021.103133

Di Falco, S., Penov, I., Aleksiev, A., van Rensburg, T.M., 2010. Agrobiodiversity, farm profits and land fragmentation: Evidence from Bulgaria. Land Use Policy 27, 763–771. https://doi.org/10.1016/j.landusepol.2009.10.007

European Commission, 2014. Commission Implementing Regulation (EU) No 809/2014 of 17 July 2014 laying down rules for the application of Regulation (EU) No 1306/2013 of the European Parliament and of the Council with regard to the integrated administration and control system, rural development measures and cross compliance, OJ L.

Geppert, C., Hass, A., Földesi, R., Donkó, B., Akter, A., Tscharntke, T., Batáry, P., 2020. Agri-environment schemes enhance pollinator richness and abundance but bumblebee reproduction depends on field size. Journal of Applied Ecology 00, 1–11. https://doi.org/10.1111/1365-2664.13682

Gocht, A., Britz, W., 2011. EU-wide farm type supply models in CAPRI—How to consistently disaggregate sector models into farm type models.

Journal of Policy Modeling, Special Issue: Impact assessment of agri-food policies in Europe:Methods, tools and applications 33, 146–167. https://doi.org/10.1016/j.jpolmod.2010.10.006

Gocht, A., Britz, W., Ciaian, P., Paloma, S.G. y, 2013. Farm Type Effects of an EU-wide Direct Payment Harmonisation. Journal of Agricultural Economics 64, 1–32. https://doi.org/10.1111/1477-9552.12005

Gocht, A., Ciaian, P., Bielza, M., Terres, J.-M., Röder, N., Himics, M., Salputra, G., 2017. EU-wide Economic and Environmental Impacts of CAP Greening with High Spatial and Farm-type Detail. Journal of Agricultural Economics 68, 651–681. https://doi.org/10.1111/1477-9552.12217

Heckelei, T., Britz, W., Zhang, Y., 2012. Positive Mathematical Programming Approaches – Recent Developments in Literature and Applied Modelling. Bio-Based Appl. Econ. 1. https://doi.org/10.13128/BAE-10567

Heinrichs, J., Kuhn, T., Pahmeyer, C., Britz, W., 2021. Economic effects of plot sizes and farm-plot distances in organic and conventional farming systems: A farm-level analysis for Germany. Agricultural Systems 187, 102992. https://doi.org/10.1016/j.agsy.2020.102992

Howitt, R.E., 1995. Positive Mathematical Programming. American Journal of Agricultural Economics 77, 329–342. https://doi.org/10.2307/1243543

ISIP, 2021. Zuckerrüben: Fruchtfolge [WWW Document]. URL https://www.isip.de/isip/servlet/isip-de/infothek/hackfruechte/zuckerrueben/fruchtfolge (accessed 6.8.21).

IT.NRW, 2019. Anbau auf dem Ackerland nach Fruchtarten - Information und Technik Nordrhein-Westfalen [WWW Document]. Landesbetrieb IT.NRW. URL https://www.it.nrw/statistik/eckdaten/anbau-auf-dem-ackerland-nach-fruchtarten-1457 (accessed 6.22.21).

Janssen, S., van Ittersum, M.K., 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agricultural Systems, Special Section: sustainable resource management and policy options for rice ecosystems 94, 622–636. https://doi.org/10.1016/j.agsy.2007.03.001

Janus, J., Markuszewska, I., 2017. Land consolidation – A great need to improve effectiveness. A case study from Poland. Land Use Policy 65, 143–153. https://doi.org/10.1016/j.landusepol.2017.03.028

Kanellopoulos, A., Reidsma, P., Wolf, J., van Ittersum, M.K., 2014. Assessing climate change and associated socio-economic scenarios for arable farming in the Netherlands: An application of benchmarking and bio-economic farm modelling. European Journal of Agronomy, Land, Climate and Resources 2020. Decision Support for Agriculture under Climate Change 52, 69–80. https://doi.org/10.1016/j.eja.2013.10.003

King, R., Burton, S., 1982. Land Fragmentation: Notes on a Fundamental Rural Spatial Problem. Progress in Human Geography 6, 475–494. https://doi.org/10.1177/030913258200600401

KTBL, 2020. Standard gross margins [WWW Document]. URL https://daten.ktbl.de/sdb/welcome.do (accessed 5.16.21).

Kuhn, Schäfer, Holm-Müller, Britz, 2019. On-farm compliance costs with the EU-Nitrates Directive: A modelling approach for specialized livestock production in northwest Germany. Agricultural Systems 173, 233–243. https://doi.org/10.1016/j.agsy.2019.02.017

Kuhn, T., Enders, A., Gaiser, T., Schäfer, D., Srivastava, A.K., Britz, W., 2020. Coupling crop and bio-economic farm modelling to evaluate the revised fertilization regulations in Germany. Agricultural Systems 177, 102687. https://doi.org/10.1016/j.agsy.2019.102687

Kuhn, T., Schäfer, D., 2018. A farm typology for North Rhine-Westphalia to assess agrienvironmental policies (Discussion Paper No. 1), Discussion Paper. Institute for Food and Resource Economics, Bonn.

Latruffe, L., Piet, L., 2014. Does land fragmentation affect farm performance? A case study from Brittany, France. Agricultural Systems 129, 68–80. https://doi.org/10.1016/j.agsy.2014.05.005

Lengers, B., 2012. Construction of different GHG accounting schemes for approximation of dairy farm emis-sions. Technical paper (ILR University of Bonn, Germany.

Lengers, B., Britz, W., Holm-Müller, K., 2014. What Drives Marginal Abatement Costs of Greenhouse Gases on Dairy Farms? A Meta-modelling Approach. Journal of Agricultural Economics 65, 579–599. https://doi.org/10.1111/1477-9552.12057

Lengers, B., Britz, W., Holm-Muller, K., 2013. Comparison of GHG-Emission Indicators for Dairy Farms with Respect to Induced Abatement Costs, Accuracy, and Feasibility. Applied Economic Perspectives and Policy 35, 451–475. https://doi.org/10.1093/aepp/ppt013

Louhichi, K., Ciaian, P., Espinosa, M., Perni, A., Gomez, S., 2015. An EU-Wide Individual Farm Model for Common Agricultural Policy Analysis. https://doi.org/10.2791/14623

Louhichi, K., Ciaian, P., Espinosa, M., Perni, A., Gomez y Paloma, S., 2018. Economic impacts of CAP greening: application of an EU-wide individual farm model for CAP analysis (IFM-CAP). European Review of Agricultural Economics 45, 205–238.

Louhichi, K., Kanellopoulos, A., Janssen, S., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Lansink, A.O., Ittersum, M.V., 2010. FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies. Agricultural Systems 103, 585–597. https://doi.org/10.1016/j.agsy.2010.06.006

Lu, H., Xie, H., He, Y., Wu, Z., Zhang, X., 2018. Assessing the impacts of land fragmentation and plot size on yields and costs: A translog production model and cost function approach. Agricultural Systems 161, 81–88. https://doi.org/10.1016/j.agsy.2018.01.001

Martini, D., 2018. Webservices auf heterogenen Datenbeständen – Methoden der Umsetzung am Beispiel der KTBL-Planungsdaten, in:

Fokus: Digitale Marktplätze und Plattformen. Presented at the GIL-Jahrestagung, Gesellschaft für Informatik e.V.

Martini, D., Herzig, D., Ladwig, G., Kunisch, M., 2014. Semantic search: finding KTBL's planning data and reusing them in IT systems. Landtechnik 69, 12–18.

M'barek, R., Barreiro-Hurlé, J., Boulanger, P., Caivano, A., Ciaian, P., Dudu, H., Goded, M.E., Fellmann, T., Ferrari, E., Paloma, S.G.Y., others, 2017. Scenar 2030-Pathways for the European agriculture and food sector beyond 2020 (Summary report). Joint Research Centre (Seville site).

Mérel, P., Bucaram, S., 2010. Exact calibration of programming models of agricultural supply against exogenous supply elasticities. European Review of Agricultural Economics 37, 395–418. https://doi.org/10.1093/erae/jbq024

Mosnier, C., Duclos, A., Agabriel, J., Gac, A., 2017. Orfee: A bio-economic model to simulate integrated and intensive management of mixed crop-livestock farms and their greenhouse gas emissions. Agricultural Systems 157, 202–215. https://doi.org/10.1016/j.agsy.2017.07.005

Mueller, L., Schindler, U., Shepherd, T.G., Ball, B.C., Smolentseva, E., Pachikin, K., Hu, C., Hennings, V., Sheudshen, A.K., Behrendt, A., Eulenstein, F., Dannowski, R., 2014. The Muencheberg Soil Quality Rating for Assessing the Quality of Global Farmland, in: Mueller, L., Saparov, A., Lischeid, G. (Eds.), Novel Measurement and Assessment Tools for Monitoring and Management of Land and Water Resources in Agricultural Landscapes of Central Asia, Environmental Science and Engineering. Springer International Publishing, Cham, pp. 235–248. https://doi.org/10.1007/978-3-319-01017-5_13

Pahmeyer, C., Britz, W., 2020. Economic opportunities of using crossbreeding and sexing in Holstein dairy herds. Journal of Dairy Science 103, 8218–8230. https://doi.org/10.3168/jds.2019-17354

Pahmeyer, C., Kuhn, T., Britz, W., 2021a. 'Fruchtfolge': A crop rotation decision support system for optimizing cropping choices with big data and spatially explicit modeling. Computers and Electronics in Agriculture 181, 105948. https://doi.org/10.1016/j.compag.2020.105948

Pahmeyer, C., Schäfer, D., Kuhn, T., Britz, W., 2021b. Data on a synthetic farm population of the German federal state of North Rhine-Westphalia. Data in Brief 36, 107007. https://doi.org/10.1016/j.dib.2021.107007

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Radulescu, C.Z., Radulescu, M., 2012. A Decision Support Tool Based on a Portfolio Selection Model for Crop Planning under Risk. Studies in Informatics and Control 21, 377–382. https://doi.org/10.24846/v21i4y201203

Reidsma, P., Janssen, S., Jansen, J., van Ittersum, M.K., 2018. On the development and use of farm models for policy impact assessment in the European Union – A review. Agricultural Systems 159, 111–125. https://doi.org/10.1016/j.agsy.2017.10.012

Schäfer, D., Britz, W., Kuhn, T., 2017. Flexible Load of Existing Biogas Plants: A Viable Option to Reduce Environmental Externalities and to Provide Demand driven Electricity? German Journal of Agricultural Economics 66.

Schroeder, L.A., Gocht, A., Britz, W., 2015. The Impact of Pillar II Funding: Validation from a Modelling and Evaluation Perspective. Journal of Agricultural Economics 66, 415–441. https://doi.org/10.1111/1477-9552.12091

Seidel, C., Britz, W., 2019. Estimating a Dual Value Function as a Meta-Model of a Detailed Dynamic Mathematical Programming Model. Bio-based and Applied Economics 8, 75–99. https://doi.org/10.13128/bae-8147

Sulewski, P., Wąs, A., Kobus, P., Pogodzińska, K., Szymańska, M., Sosulski, T., 2020. Farmers' Attitudes towards Risk—An Empirical

Study from Poland. Agronomy 10, 1555. https://doi.org/10.3390/agronomy10101555

van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., van der Wal, T., Wien, J.E., Wolf, J., 2008. Integrated assessment of agricultural systems – A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150–165. https://doi.org/10.1016/j.agsy.2007.07.009

Chapter 6
Conclusion

The overall research aim of this thesis is to show potential ways for modeling highly detailed policies and technologies in BEFMs used for policy and technology analysis as well as for decision support. The following chapter summarizes the contributions of the previous chapters in the thesis. In addition, the methodologies applied in the different chapters are critically analysed, and a research outlook is provided.

6.1 Major contributions of the thesis

The major contributions of the thesis are three-fold: First, the thesis presents practical examples for modeling highly detailed technology innovations, policy regulations, and DSS that may serve as a basis for fellow researchers seeking to study similar topics in other regions or contexts. Second, it provides a general methodology for deriving a highly detailed synthetic farm population together with an actual dataset for the German federal state of North Rhine-Westphalia. Third, the thesis explores how different levels of detail regarding the depiction of crop choices influence the results of a BEFM, and gives recommendations for which types of use-cases can benefit from higher levels of detail.

As an example for modeling a highly detailed technology innovation in a BEFM, the thesis analyses the economically optimal use of sex sorted semen (called "sexed semen") and crossbreeding among a dairy farm population from the German federal state of North Rhine-Westphalia. Given the high level of detail in the implementation of the technology, the analysis reveals that the profit maximizing sexed semen and crossbreeding utilization is highly heterogeneous among the farms in the study area. Farms endowed with stocking densities lower than 2 LU/ha are in general found to produce excess heifers for sale, whereas farms endowed with stocking densities of more than 2 LU/ha are found to be producing crossbred calves and using sexed semen solely to produce replacement animals. Approximately 25.3% of all inseminations are found to involve female sexed dairy semen. Furthermore, beef semen (both sexed and conventional) for producing crossbred calves is found to be used in approximately 21.5% of the inseminations. Combining sexed semen with crossbreeding is found to increase profits from 0 €/cow/year to 568 €/cow/year, with an average of 79.42 €/cow/year. As a result of using sexed semen, farms with stocking densities below 2 LU/ha and above average replacement rates of 40% are found to showcase higher profit increases as a result of selling more excess heifers. Crossbreeding adoption and overall sexed semen adoption are most affected by stocking density and (farm) average cow longevity, as well as additional costs for sexed semen and sexed semen accuracy. These results demonstrate that modern breeding technologies have the potential to improve dairy farm profits, although they must be viewed in the context of farm-specific production settings. For policy makers, it is interesting to note that sexed semen can mitigate the issue of male calves being born that are of almost no economic value. In the context of an ongoing debate about animal welfare, the technology may help to reduce the socio-ethical concerns raised by the production of male dairy calves (Balzani et al., 2021). Furthermore, the intensification of beef production originating from dairy systems has been shown to produce less greenhouse gas emissions per unit of product compared to traditional suckler systems, thus highlighting the potential of the combination of sexed semen and crossbreeding (Hietala et al., 2014).

As an example for modeling a policy with measures targeting single fields of a farm in a BEFM, the thesis presents a novel DSS called ‘Fruchtfolge’ assisting farmers with finding a cost minimal adoption strategy to the newly revised FO (revision of 2020) in Germany. The DSS shows a potential way for providing decision makers with a crop and management recommendation for each of their fields based on the solution of a BEFM. In order to provide recommendations in the level of detail required by the policy, the BEFM accounts for field specific location factors, labour endowments, field-to-farm distances and relevant policy measures. With the ‘Fruchtfolge’ DSS, farm, location and management characteristics are automatically incorporated into a user-friendly tool. Users are provided with instant feedback about alternative management options, and may generally create a first economically optimal cropping plan in less than five minutes. In a case study application involving a farm managing fields both outside and inside of a nitrate sensitive area according to the revised FO, the DSS is shown to mitigate the farms compliance costs to the revised FO by more than 5%. These results demonstrate that the ‘Fruchtfolge’ DSS presented in the thesis can aid decision makers in optimizing cropping choices in complex environments and reduce potential losses of profits. For policy makers it is interesting to note that the ‘Fruchtfolge’ DSS can also help to reduce bureaucratic obstacles, as the DSS automatically provides the fertilizer planning sheets required by the FO 2020 without any manual data entry. As a result, using the ‘Fruchtfolge’ DSS which provides the fertilizer planning sheets as well as an optimized cropping and fertilization plan is arguably faster than filling out the respective sheets manually. Furthermore, as the DSS provides farmers with the most cost-effective adoption strategy towards the regulations of the new FO, and will also push notifications when farmers select management options that are non-compliant, they are nudged to comply with the legislation which may help to reduce negative externalities otherwise caused by violations.

As previously stated, the thesis also provides data and a methodology for creating a synthetic farm population of the German federal state of North Rhine-Westphalia. The synthetic farm population outlined in the thesis is largely based on the German Farm Structure Survey 2016, as well as plot specific crop data from 2019/2020. The underlying farm population is derived from a farm typology at administrative unit level to which observed plots are allocated after the initial farm assignment. The resulting dataset contains 25,858 farms and covers 1.3 million ha of agricultural land, provided at single plot scale in multiple digital formats. The single plot data includes information regarding the managing farm including the randomly assigned farm location, the number of livestock the farm is endowed with, the cultivated crop on the plot, as well as the corresponding administration units. In addition, relevant spatial data such as yield information, soil characteristics, as well as monitoring data on environmental status are included. By making use of the data provided by the synthetic farm population, a variety of analyses involving farm, agent-based, and bio-physical models can be carried out. The dataset can also serve as a test dataset for a variety of farm models requiring spatially explicit information. Due to its general design, the methodology for creating the farm population can be transferred to other regions where access to individual farm data records is restricted.

With regards to the methodological contributions of the thesis, a thorough assessment of the prevalent ways for modeling crop choices in BEFMs, as well as their influence on simulation results is given. In the thesis, three possible approaches for modeling crop choices motivated from the literature are considered: 'single plots' with one crop per season, crop shares of land differentiated by soil type, called 'categorized', and crop shares on all arable land, termed 'aggregate'. The analysis is conducted using the previously mentioned, highly detailed synthetic farm population from North Rhine-Westphalia. The results of the comparative analysis indicate that the 'aggregate' and 'categorized' land endowment approaches produce similar simulation results, which however diverge from the 'single plot' approach. The results indicate that crop choices per farm differ by approximately 11% between the spatially explicit 'single plot' and the 'aggregate' land endowment approach in the case study region of North Rhine-Westphalia. Total work requirements are found to be on average 10% higher in the 'aggregate' approach compared to the 'single plot' approach, while energy requirements are relatively similar (average difference of 2.2%). Among other factors, the results indicate that this difference is highly correlated with the number of plots a farm is endowed with. For instance, the average difference in crop choices increases from the sample average of 11% to 20.8% for those farms that are endowed with less than 10 plots (~ 50% of the case study population). Differences in simulated farm profits when comparing the 'aggregate' land endowment approach to the 'single plot' approach are found to range between -306 €/ha to 434 €/ha. The results suggest that for analyses using BEFMs with a focus on aggregate results over a larger sample of farms, both the 'aggregate' and 'categorized' land endowment approaches are sufficiently accurate in case of similar average numbers of plots per farm as in our study. If single farm results or variability in the population are targeted, we propose to incorporate the 'single plot' approach in bio-economic farm analyses. The same holds for DSS focusing on individual farm responses to policy changes or technology adoption. For policy makers it is interesting to note that solely BEFMs operating on the single plot level can accurately depict restrictions targeting individual plots, such as the fertilization management restrictions imposed by the FO 2020 enacted in so-called 'red areas'. Studies capturing the potential environmental benefits of agri-environmental measures could also profit from the depiction of individual plots, potentially including further spatial dimensions such as proximity to forests or waterways.

6.2 Methodological discussion and research outlook

The dissertations methodological focus primarily lies in the development, extension, and application of BEFMs used for policy and technology evaluation as well as decision support. BEFMs are especially suited for ex-ante policy and technology evaluation, and have a few advantages over other methods used for impact assessments as outlined by Jannsen and van Ittersum (2007): First, BEFMs are based on a constrained optimization procedure, and thereby reflect farmers situation striving to improve their operation with limited resources. By design, the BEFMs FarmDyn and 'Fruchtfolge' follow different approaches for the specification of a farm's available resources. In FarmDyn, available resources are specified exogenously based on data collected from case study farms, farm typologies (Kuhn and Schäfer, 2018) or the synthetic farm population outlined in Chapter 4. 'Fruchtfolge' on the other hand follows a participatory approach, presenting assumptions with regards to farm planning data as a default that can interactively be adjusted towards the situations in which the decision makers find themselves. Given the available resources, both models proceed to find the economically optimal resource allocation with respect to the given prices, yields, and costs.

Second, BEFMs can simultaneously consider multiple possible production activities, restrictions, new production techniques, and may include (indirect) linkages between crop and livestock production. These linkages are especially well depicted in the holistic BEFM FarmDyn, as it allows for the specification of multiple farm branches including livestock. Considering the inclusion of novel technologies (see Chapter 2), these linkages allow for analysing research questions that involve a competition among intra-farm resources across farm branches. As the 'Fruchtfolge' BEFM does not incorporate livestock or other farm branches, the intra-farm competition for resources can only be considered among different crops and their associated management options.

Third, BEFMs allow for an analysis of changing parameter values through sensitivity analysis. In the context of the thesis, sensitivity analysis is primarily used to identify factors that are most influential for the economic viability of a technology (Chapter 2), and to assess which factors influence the difference in BEFM simulation results induced by the different depictions of cropping choices Chapter 5). While a framework for a structured sensitivity analysis using Latin Hypercube Sampling (LHS) is an integral part of the FarmDyn model (Chapter 2), the 'Fruchtfolge' DSS does not incorporate a pre-defined tool for such an assessment. Since the 'Fruchtfolge' DSS follows a participatory approach, users may interactively adjust parameters in the user-interface and immediately observe the effects of the changes on the results. However, as the BEFM backing the 'Fruchtfolge' DSS can also be utilized without its user-interface, a systematic sensitivity analysis can be performed as demonstrated in Chapter 5.

Fourth, BEFMs may in general be used for both short-term predictions and long-term trend analysis. While FarmDyn supports fully (recursive) dynamic simulations, the thesis solely makes use of the comparative static version of the model as investment decisions are not specifically targeted in the analysis. Similarly, also the 'Fruchtfolge' DSS deliberately focuses on a planning horizon of a single year, indirectly including long-term effects through the inclusion of maximum crop shares and crop waiting periods. As pointed out by Jannsen and van Ittersum (2007), the long-term predictive power of normative mechanistic BEFMs such as FarmDyn and 'Fruchtfolge' is restricted, rendering the rather short term analysis as done in this thesis more appropriate for this particular type of BEFM.

Furthermore, as noted by Blanco (2016), farm-level analyses using BEFMs allow for greater modeling flexibility for capturing farm heterogeneity, environmental effects, and economic performance than for instance other types of mathematical programming models such as partial and general equilibrium models. In the context of the underlying thesis, these methodological advantages facilitate some of the major findings and underline their scientific contributions.

With regards to the technology evaluation study presented in Chapter 2, applying the highly detailed BEFM FarmDyn for assessing the impacts of novel breeding technologies such as sexed semen poses few advantages over process-based approaches frequently employed in the literature. As previously stated, the FarmDyn BEFM endogenously optimizes multiple decision variables simultaneously. These variables include herd entry and exit dates, fodder production and use of concentrates, grassland management, manure storage and management, allocation of labour to cash crops and herd management, animal housing and machinery utilization, as well as inputs required for crop production. Opposed to a static simulation approach where the levels of few decision variables such as feed uses, heifer breeding strategies, and crop allocation are mostly pre-determined, using a BEFM such as FarmDyn allows for exploring the profit-maximizing levels of these variables for each farm. The resulting profit-maximizing strategies for crossbreeding and sexed semen use are therefore demonstrating the full potential of the available options since the BEFM incorporates the complexity of the decision at the whole farm level.

Considering the heterogeneity of the results regarding the economically optimal sexed semen and crossbreeding usage among the farms, the approach for studying technology innovations outlined in the thesis highlights the importance of studying the whole variety of farms in a population. In combination with the large scale sensitivity analysis conducted with regards to the various bio-economic parameters influencing the adoption decision (also called meta-modeling in this context (see Kuhn et al., 2019; Lengers et al., 2014; Seidel and Britz, 2019)), the given approach can be seen as a general methodology to estimate potential economic returns for a technology for which technological parameters and related costs and benefits cannot be derived from real farm observations yet.

Despite these advantages, BEFMs also suffer from shortcomings. By design, the BEFMs considered in this thesis solely depict the supply side, thus treating market dependant figures such as input and output prices as exogenous factors. By disregarding potential market effects resulting from the supplier's decision, potential feedback loops of the supply side production decision on input and output prices cannot be captured by the BEFMs as used in this thesis. Considering the concrete example of sexed semen and crossbreeding uptake presented in Chapter 2, feedback of the simulated supply increase of crossbred animals on the producer price of crossbred calves as discussed by De Vries et al. (2008) could be expected. As stated earlier, market models such as (partial) equilibrium models incorporate such market feedback by design, however miss the detailed depiction of technical production processes required by technology evaluation studies such as the one outlined in Chapter 2.

Another shortcoming is the fact that the BEFMs used in this thesis do not consider technology diffusion as a factor, despite its known implications on the adoption of novel technologies (Barbuto et al., 2019). As a result, the BEFMs may instantaneously switch to the novel technology upon availability in the model. However, as the thesis does not aim to predict actual responses from farmers to a novel policy or technology, but rather show promising and economically feasible alternatives, this shortcoming does not affect the results presented in the thesis.

A further issue of mechanistic BEFMs as used in this thesis is their tendency of overspecialization in production decisions (Gocht et al., 2016; Janssen and van Ittersum, 2007). By design, these (mixed-integer) LP models will select the most profitable production option until all required resources are exhausted or another constraint becomes binding. To give a concrete example in the context of the thesis, a (mixed integer) LP model will choose sexed semen over conventional semen for artificial insemination of a dairy cow if its use increases the farms profits at least marginally. Considering no other constraints limiting sexed semen uptake, the LP will subsequently select sexed semen for all dairy cows of the farm. In order to limit the overspecialization behavior of an LP, the PMP approach can be applied in situations were observed levels of decision variables are available (Heckelei et al., 2012; Howitt, 1995). However, as the focus of the underlying thesis rather lies in the exploration of feasible, profit maximizing production alternatives, and not in the prediction of future uptake values, this limitation does again not apply to the results outlined in this thesis.

Another criticism regarding BEFMs as used in this thesis is their behavioral assumption of a fully informed, rational, profit maximizing decision maker (Malek et al., 2019; Reidsma et al., 2018). Many studies suggest that, despite being an important factor, economic rationale is only one consideration out of many that influence the decision process (An, 2012; Levine et al., 2015; Malek et al., 2019; Nualnoom et al., 2016). In response to this limitation, alternative model types such as agent-based models (ABM) have been developed (Huber et al., 2021; Seidel and Britz, 2019). Opposed to traditional LPs as used in this thesis, in an ABM each decision maker presents an agent that can interact with surrounding agents. By including these agents, ABMs allow for the inclusion of behavioral factors as well as cognitive, emotional, personal and social processes, thus depicting a more realistic decision process compared to simple LPs (Huber et al., 2021). As the main purpose of ABMs again lies in improving the quality of studies aiming to predict farmers behavior, the criticism is less relevant with regards to the BEFMs used in the underlying thesis exploring economically viable production alternatives. As a final remark, it has to be acknowledged that throughout the thesis deterministic BEFMs have been used[1]. In the deterministic models, no parameters are stochastic, and therefore also no variables are state-contingent. Therefore, volatility in yields as well as input and output prices are not endogenously considered in the BEFMs used in the thesis. While a large-scale sensitivity analysis such as the one conducted in

[1] While the BEFM FarmDyn includes a module covering several risk models such as value at risk (VaR), MOTAD (Hazell, 1971), and Target MOTAD (Tauer, 1983), the risk module has not been used in the underlying thesis.

Chapter 2 generally captures the overall volatility of the parameter ranges in the deterministic model, it cannot endogenously reflect the associated variance in farm income. Using a deterministic BEFM with the assumption of risk neutrality poses a few issues with regard to the selection of the (mathematically) optimal activities, as these may not reflect the utility maximizing activities a real farmer would prefer (Musshoff and Hirschauer, 2007). Among others, Sulewski et al. (2020) show that farmers are rather risk-averse than risk-neutral. Subsequently, they prefer to reduce the (semi-) variance in their expected farm income when maximizing their utility (Rosa et al., 2019). The results of a deterministic BEFM assuming risk-neutrality might therefore showcase higher income variance than acceptable by the decision maker, thus rendering the solution not optimal considering the decision makers risk preferences.

While again this limitation affects exploratory studies such as the one outlined in Chapter 2 only to a lesser extent, it does affect BEFMs used for decision support such as the 'Fruchtfolge' DSS presented in Chapter 3. In the DSS, only expected yields and prices are considered in the optimization, without considering their inherent variances. Depending on the user's attitude towards risk, the presented solution might therefore exceed the income variance threshold accepted by the user. However, in order to incorporate the user's attitude towards risk in the DSS, the user's individual risk aversion coefficient would need to be determined. Determining farmers' risk aversion coefficients has been proven to be a challenging task (Cao et al., 2011; Sulewski et al., 2020). Musshoff and Hirschauer (2007) present a practical approach for determining the willingness to accept risks by taking the variance of the total gross margin inherent to the crop choices and activities chosen by the farmer in previous years. This empirically observed variance in total gross margin is then set as a constraint when maximizing the expected total gross margin of the farm. As the previously selected cropping choices by the farmer, as well as price and yield variations are already present in the 'Fruchtfolge' DSS, this approach could be incorporated in future versions of the DSS.

Besides these limitations, novel aspects of the BEFMs developed in this thesis also contribute to the literature in multiple aspects. Especially with regards to the use of BEFMs in DSS as outlined in Chapter 3, few methodological concepts can be mentioned. For instance, the ‘Fruchtfolge’ DSS is among the first BEFMs aimed at optimizing cropping choices to follow principles of user-centered design as outlined by Rose (2017, 2016). Among others, the user-centered design aspects relate to fully automated data collection solely requiring users to provide their CRN to access EU direct payment applications of their farm for an initial optimization. In the context of the thesis, this functionality has been made publicly available in a separate open-source software repository[2]. As highlighted by Britz (2014), adding a user-interface (the DSS) on top of a BEFM also eases the communication between the BEFM and the user, shielding away details of the underlying economic programming model. The user-friendly and visually attractive interface of the ‘Fruchtfolge’ DSS can therefore help to communicate the results of the optimization in a more efficient manner. Furthermore, the user-interface of the ‘Fruchtfolge’ BEFM allows for quickly exploring alternative cropping and management choices as well as their consequences on total gross margin and fertilization strategies. This interactive approach to presenting the model results to users has already been proposed by McCown (2001), improving dialogue between the model and the users and thus building trust and confidence in the results. In order to provide users with meaningful error messages and assistance in case their farm specification turns the model mathematically infeasible, slack variables are introduced for each decision variable in the underlying BEFM. Also, as the DSS is provided as a web application, it enables continuous updates by the provider, and flexible access for decision makers. However, a thorough assessment of whether these novel approaches will help at overcoming the often-observed underuse of DSS at farm-scale (Rose et al., 2016) is still missing. Future research should therefore test whether decision makers are willing to accept and use the DSS, and find potential room for improvement. Also, potential integrations of the ‘Fruchtfolge’ DSS with other farm

[2] The open-source software package *harmonie* aimed at harmonizing the various EU direct payment application files can be found under following software versioning repository: https://github.com/fruchtfolge/harmonie

management tools could be imagined, allowing for a better alignment to the established planning routines of farmers and their advisers.

The synthetic farm population presented in Chapter 4 is largely based on a farm typology presented by Kuhn and Schäfer (2018) which in turn is based on official statistical data from the Farm Structure Survey 2016, and the Census of Agriculture. While these data sources cover a majority of the farms and their farm characteristics, they express a few shortcomings that affect the quality and reliability of the resulting synthetic farm population. First, the distribution of farm characteristics queried for the creation of the original farm typology excludes few specific farm types. Especially farms specialized in vegetable or perennial crop production are not explicitly covered in the dataset, and are therefore also not part of the synthetic farm population. Due to these missing farms, the synthetic farm population does not cover all of the farms in the federal state of North Rhine-Westphalia, but rather 77% of the farms cultivating 89% of the agricultural area. Since the farm frequency tables provided at LAU level include all of the farms instead, some farms presented in the frequency table cannot be matched to a corresponding farm from the Farm Structure Survey dataset. Second, farms that are split into multiple smaller legal units, but managed as a single farm in practice cannot be detected in the dataset. As a result, there may be more farms in the population than actual farms operating in reality. Accordingly, the farms in the synthetic farm population may expose smaller endowments and farm characteristics than their real counterparts. When using the synthetic farm population in policy or technology evaluation studies, the share of smaller farms within the farm population may be overestimated. Resulting from these limitations, a potential update to the synthetic farm population should therefore incorporate all farm types present in the recently released Farm Structure Survey 2020. Furthermore, it should be extended using surveys and expert knowledge in order to mitigate the issue of single farms being split into smaller units as present in the synthetic population.

The comparison of the possible approaches for depicting cropping decisions in BEFMs outlined in Chapter 5 highlights the influence of the chosen approach on the simulation results. The results of the analysis suggest two distinct effects that influence the difference between the three identified cropping choice approaches. On the one hand, the 'aggregate' land endowment approach, treating a farm's land endowment as a homogenous land mass with equal characteristics (e.g. soil quality, farm-field-distance, size), may suffer from an aggregation bias when compared to the spatially explicit 'single plot' approach. On the other hand, the assumption of indivisibilities of plots in the 'single plot' approach introduces an 'indivisibility bias' compared to the 'aggregate' and 'categorized' approach. Aggregation bias has been acknowledged as a more serious problem in BEFMs used for sectoral analysis for decades (Buckwell and Hazell, 1972; Day, 1963; Önal and McCarl, 1991; Paris and Rausser, 1973). It has to be mentioned that the term "aggregation bias" as coined by the literature generally discusses the bias introduced by aggregating farms within a region into a single farm cluster, which is in turn solved by a BEFM. However, the aggregation bias as dealt with in Chapter 5 rather specifies the bias introduced by aggregating a single farms plots into a single land mass, thus being an intra-farm- opposed to an inter-farm bias. In this context, formulating conditions potentially reducing the aggregation bias to zero (as done by Paris and Rausser (1973)) is far from trivial, as each of the possible cropping options react differently to changes in plot characteristics such as plot size, distance to the farmstead, and soil quality. While previously, a lack of data on the single farm level made aggregating farms into a single farm cluster necessary, the provision of highly detailed farm populations (as described in Chapter 4) allows for such micro level analysis as the one conducted in Chapter 5. Opposed to the aggregation bias, the indivisibility bias ascribed to the assumption of plots being indivisible can be considered as more debatable though. In the context of the thesis, the definition of a plot follows Nesme et al. (2010), stating that a plot is the smallest homogeneously managed unit a farmer operates on. Since the plots contained in the synthetic farm population are the actual plots (including their geometries and characteristics) farmers have used in their applications for the EU direct payments, it is assumed that these plots already present practically optimal, manageable units of land which should not be split into smaller units. However, when comparing the plots entered into the EU direct payment application program for a few years[3], it can be observed that farmers indeed split some of their fields into smaller units from time to time. Future research should therefore elicit if the 'single plot' approach using binary variables is more appropriate in order to correctly depict the crop planning decision problem, or whether a highly detailed 'categorized' approach using continuous variables is better suited.

The methodologies applied in this thesis can contribute to both ex-ante and ex-post analysis on the economic impact of novel policies and technologies on farms. Policy initiatives such as the European Green Deal, the upcoming revision of the Common Agricultural Policy, as well as novel technologies as for instance agricultural robots pose several potential areas of applications for the BEFMs presented in the context of the thesis. Furthermore, the extended use of DSS may help to disseminate insights gathered by the use of BEFMs to farmers and farm advisers.

---

[3] The official datasets can be obtained from IT.NRW (2021)

6.3 References

An, L., 2012. Modeling human decisions in coupled human and natural systems: Review of agent-based models. Ecological Modelling, Modeling Human Decisions 229, 25–36. https://doi.org/10.1016/j.ecolmodel.2011.07.010

Balzani, A., Aparacida Vaz do Amaral, C., Hanlon, A., 2021. A Perspective on the Use of Sexed Semen to Reduce the Number of Surplus Male Dairy Calves in Ireland: A Pilot Study. Front Vet Sci 7, 623128. https://doi.org/10.3389/fvets.2020.623128

Barbuto, A., Lopolito, A., Santeramo, F.G., 2019. Improving diffusion in agriculture: an agent-based model to find the predictors for efficient early adopters. Agricultural and Food Economics 7, 1. https://doi.org/10.1186/s40100-019-0121-0

Blanco, M., 2016. Policy impact assessment., in: Shrestha, S., Barnes, A., Ahmadi, B.V. (Eds.), Farm-Level Modelling: Techniques, Applications and Policy. CABI, Wallingford, pp. 1–13. https://doi.org/10.1079/9781780644288.0001

Britz, W., 2014. A New Graphical User Interface Generator for Economic Models and its Comparison to Existing Approaches [WWW Document]. German Journal of Agricultural Economics. https://doi.org/ISSN 0002-1121

Buckwell, A.E., Hazell, P.B.R., 1972. Implications of Aggregation Bias for the Construction of Static and Dynamic Linear Programming Supply Models*. Journal of Agricultural Economics 23, 119–134. https://doi.org/10.1111/j.1477-9552.1972.tb01436.x

Cao, R., Carpentier, A., Gohin, A., 2011. Measuring farmers' risk aversion: the unknown properties of the value function (No. 114623), 2011 International Congress, August 30-September 2, 2011, Zurich, Switzerland, 2011 International Congress, August 30-September 2, 2011, Zurich, Switzerland. European Association of Agricultural Economists.

Day, R.H., 1963. On Aggregating Linear Programming Models of Production. Journal of Farm Economics 45, 797–813. https://doi.org/10.2307/1235749

De Vries, A., Overton, M., Fetrow, J., Leslie, K., Eicker, S., Rogers, G., 2008. Exploring the Impact of Sexed Semen on the Structure of the Dairy Industry. Journal of Dairy Science 91, 847–856. https://doi.org/10.3168/jds.2007-0536

Gocht, A., Ciaian, P., Espinosa, M., Gomez y Paloma, S., 2016. Linkage of a farm group model to a partial equilibrium model. Farm-level modelling: techniques, applications and policy 189–205.

Heckelei, T., Britz, W., Zhang, Y., 2012. Positive Mathematical Programming Approaches – Recent Developments in Literature and Applied Modelling. Bio-Based Appl. Econ. 1. https://doi.org/10.13128/BAE-10567

Hietala, P., Bouquet, P., Juga, J., 2014. Effect of replacement rate, crossbreeding and sexed semen on the efficiency of beef production from dairy herds in Finland. Acta Agriculturae Scandinavica, Section A — Animal Science 64, 199–209. https://doi.org/10.1080/09064702.2015.1013978

Howitt, R.E., 1995. Positive Mathematical Programming. American Journal of Agricultural Economics 77, 329–342. https://doi.org/10.2307/1243543

Huber, R., Xiong, H., Keller, K., Finger, R., 2021. Bridging behavioural factors and standard bio-economic modelling in an agent-based modelling framework. Journal of Agricultural Economics. https://doi.org/10.1111/1477-9552.12447

IT.NRW, 2021. Landwirtschaftlich relevante Daten [WWW Document]. URL https://www.opengeodata.nrw.de/produkte/umwelt_klima/bodennutzung/landwirtschaft/ (accessed 10.4.21).

Janssen, S., van Ittersum, M.K., 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models.

Agricultural Systems 94, 622–636. https://doi.org/10.1016/j.agsy.2007.03.001

Kuhn, T., Schäfer, D., 2018. A farm typology for North Rhine-Westphalia to assess agrienvironmental policies (Discussion Paper No. 1), Discussion Paper Series. Institute for Food and Resource Economics, Bonn.

Kuhn, T., Schäfer, D., Holm-Müller, K., Britz, W., 2019. On-farm compliance costs with the EU-Nitrates Directive: A modelling approach for specialized livestock production in northwest Germany. Agricultural Systems 173, 233–243. https://doi.org/10.1016/j.agsy.2019.02.017

Lengers, B., Britz, W., Holm-Müller, K., 2014. What Drives Marginal Abatement Costs of Greenhouse Gases on Dairy Farms? A Meta-modelling Approach. Journal of Agricultural Economics 65, 579–599. https://doi.org/10.1111/1477-9552.12057

Levine, J., Chan, K.M.A., Satterfield, T., 2015. From rational actor to efficient complexity manager: Exorcising the ghost of Homo economicus with a unified synthesis of cognition research. Ecological Economics 114, 22–32.

Malek, Ž., Douw, B., Vliet, J.V., Zanden, E.H.V.D., Verburg, P.H., 2019. Local land-use decision-making in a global context. Environ. Res. Lett. 14, 083006. https://doi.org/10.1088/1748-9326/ab309e

McCown, R.L., 2001. Learning to bridge the gap between science-based decision support and the practice of farming: Evolution in paradigms of model-based research and intervention from design to dialogue. Aust. J. Agric. Res. 52, 549–572. https://doi.org/10.1071/ar00119

Musshoff, O., Hirschauer, N., 2007. What benefits are to be derived from improved farm program planning approaches? – The role of time series models and stochastic optimization. Agricultural Systems 95, 11–27. https://doi.org/10.1016/j.agsy.2007.03.007

Nesme, T., Lescourret, F., Bellon, S., Habib, R., 2010. Is the plot concept an obstacle in agricultural sciences? A review focussing on fruit production. Agriculture, Ecosystems & Environment 138, 133–138. https://doi.org/10.1016/j.agee.2010.04.014

Nualnoom, P., Wehrmeyer, W., Morse, S., 2016. Analysing household decision-making on oil palm cultivation in Thailand. Journal of Land Use Science 11, 560–578. https://doi.org/10.1080/1747423X.2016.1204019

Önal, H., McCarl, B.A., 1991. Exact Aggregation in Mathematical Programming Sector Models. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 39, 319–334. https://doi.org/10.1111/j.1744-7976.1991.tb03575.x

Paris, Q., Rausser, G.C., 1973. Sufficient Conditions for Aggregation of Linear Programming Models. American Journal of Agricultural Economics 55, 659–666. https://doi.org/10.2307/1238355

Reidsma, P., Janssen, S., Jansen, J., van Ittersum, M.K., 2018. On the development and use of farm models for policy impact assessment in the European Union – A review. Agricultural Systems 159, 111–125. https://doi.org/10.1016/j.agsy.2017.10.012

Rosa, F., Taverna, M., Nassivera, F., Iseppi, L., 2019. Farm/crop portfolio simulations under variable risk: a case study from Italy. Agricultural and Food Economics 7, 8. https://doi.org/10.1186/s40100-019-0127-7

Rose, D.C., Parker, C., Fodey, J., Park, C., Sutherland, W.J., Dicks, L.V., 2017. Involving stakeholders in agricultural decision support systems: Improving user-centred design. International Journal of Agricultural Management 6, 10. https://doi.org/10.5836/ijam/2017-06-80

Rose, D.C., Sutherland, W.J., Parker, C., Lobley, M., Winter, M., Morris, C., Twining, S., Ffoulkes, C., Amano, T., Dicks, L.V., 2016. Decision support tools for agriculture: Towards effective design and delivery. Agricultural Systems 149, 165–174. https://doi.org/10.1016/j.agsy.2016.09.009

Seidel, C., Britz, W., 2019. Estimating a Dual Value Function as a Meta-Model of a Detailed Dynamic Mathematical Programming Model. Bio-based and Applied Economics 8, 75–99. https://doi.org/10.13128/bae-8147

Sulewski, P., Wąs, A., Kobus, P., Pogodzińska, K., Szymańska, M., Sosulski, T., 2020. Farmers' Attitudes towards Risk—An Empirical Study from Poland. Agronomy 10, 1555. https://doi.org/10.3390/agronomy10101555

Chapter 7 Appendix

7.1 Appendix Chapter 2

All herds (h, being cows, heifers, calves) are differentiated by age and sex (elements of the set h), breeds (b, being Holstein and the beef breed), feeding regime (f) and month (m) in a year (t), which is fixed to a single year in the underlying analysis.

Each animal in a herd is described by the same herd average characteristics for values such as

The current size of a standing herd α summed up over all feeding regimes is defined as the number of animals that enter the herd at the beginning of the production process N, subtracted by the leaving ones K (see Equation 1).

$$\sum_{f} \alpha_{h,b,f,m,t} = N_{h,b,m,t} - K_{h,b,m,t}$$

A production process (and the corresponding herd) may refer to a fattening phase or a performance group. The heifers are differentiated by a minimum of two production processes, starting at the final weight of a female calf and ending at the starting weight of the young dairy cows. This way, daily weight gains, first calving ages, and the corresponding feeding requirements can be modeled in great detail. Additional processes may be added to reflect different first calving ages or differences in weight gains.

In order to maintain a steady herd balance, new animals need to enter a production process. The number of new animals joining a production process reflects the level of suitable delivery processes (D) which refer to the previous fattening/growing phase (in the case of heifers) or replacements (for cows). Also, new calves serve as a delivery process for either the first stage phase of heifers or bulls.

The number of animals joining a herd at the beginning of a production process (N) is described as follows:

$$N_{h,b,m,t} = D_{h,b,m,t} + B_{h,b,m,t} - S_{h,b,m,t} - \sum_{h1} N_{h1,b,m,t}$$

The number of animals stemming from a suitable delivery process (D), including animals bought to the herd from the market (B), minus animals sold to the market (S), and the sum of other production processes competing for the same delivery process ($h1 \subseteq h$). This concept can be illustrated by the following example: Suppose a female calf is born on the farm which might be used in a heifer process for the own dairy replacements or be sold after an initial rearing period of three weeks. The amount of animals that enter the farm's heifer growing process (N) is then equal to the newborn calf (= 1), subtracted by the calves sold to the market (= 0 or 1), ignoring any additional animals bought from the market. Through this approach, flexibility in the herd dynamics is ensured to simulate economic optimal decisions.

In the comparative static setting, the number of animals entering a production process is additionally corrected by the production length of the process ($l = \frac{1}{\text{production length}}$). This way, different amounts of lactations and lactation lengths that affect cow replacement rates are mapped correctly to an average year. The adjusted formula characterizing the number of animals entering a production process can then be specified as follows:

$$N_{h,b,m,t} = \frac{D_{h,b,m,t} + B_{h,b,m,t} - S_{h,b,m,t} - \sum_{h1} N_{h1,b,m,t} \cdot l_{h1,b,m,t}}{l_{h,b,m,t}}$$

The amount of calves being born in a month is defined by the sum of cows entering the cow herd throughout the simulation period multiplied with the probability of having a calf ($\widehat{p}_m$) in that particular month:

$$D_{calves,b,m,t} = \sum_{\text{cows,t1,m1,mDist}} N_{cows,b,m1,t1} \cdot \widehat{p}_m\,(\text{cows}, \text{mDist}) \quad \forall\ \text{mDist} = \Delta\text{mDist}_{t,m,t1,m1}$$

where

$$\begin{aligned}
\text{calves, cows} &\subseteq h \\
t1 &\equiv t \\
m1 &\equiv m \\
\text{mDist} &= \text{Consecutive months in the simulation period} \\
\Delta\text{mDist}_{t,m,t1,m1} &= \text{Difference in months between year t, month m, and year t1, month m1}
\end{aligned}$$

The calving distribution is largely dependent on the average calving interval (μ) and amount of inseminations needed for conception (IE, insemination effort). The calving interval is defined by the calving to conception period plus the duration of the gestation. Depending on the required insemination effort (IE), the calving to conception period (CC) may be equal to the calving to first service period (CFS) or additionally prolonged by the additional service interval (ASI). If the IE is equal to one, the cow conceives at the first insemination, and the ASI is equal to zero. Any additional insemination is assumed to increase the ASI by 30 days, respectively (Römer et al., 2013).For a given average calving distribution (μ) and average insemination effort (AIE), the calving to first service period of a herd is calculated as

$$\text{CFS} = \mu - (30 \cdot (\text{AIE} - 1))$$

A CI of 408 days (~13.38 months), and a herd average IE of 2.3 (Römer et al., 2013; Volkmann et al., 2014) subsequently result in a constant CFS of 88 days for our given example herd. Consequently, the calving distribution of a single animal is solely driven by its individual IE in this simplified model. The maximum IE is assumed to be 3.6 (Römer et al., 2013).

The above excursus highlights the importance of the insemination effort on the calving interval and thus the calving distribution of the herd. In order to depict the resulting uncertainty in the model FarmDyn, a maximum entropy estimator is used for the estimation of calving probabilities for a particular month ($\widehat{p}_m$). Given an average calving interval of a herd (μ), the a priori information assumes that calving is equally likely between the two months $m_{\text{lo}} = \lfloor\frac{\mu}{30.5}\rfloor$ and $m_{\text{up}} = \lceil\frac{\mu}{30.5}\rceil$ surrounding the average calving interval. Therefore, a calving interval of 408 days (~13.38 months) as in our example would assume a priori equal likelihood of calvings in the months $m_{\text{lo}} = 13$

and $m_{\text{up}} = 14$. The actual average interval is then recovered by forcing the sum of the probabilities $\widehat{p}_m$ multiplied with the corresponding month (here either month 13 or 14) to be equal to the average calving interval μ, maximizing the entropy which can be interpreted as picking the posteriori probabilities closest to the a priori ones:

$$\max \widehat{H} \quad = -\sum_{m}^{N} \widehat{p}_m \ln \frac{\widehat{p}_m}{N}$$

subject to

$$\sum_{m}^{N} \widehat{p}_m \quad = 1$$

$$\sum_{m}^{N} \widehat{p}_m \cdot m \quad = \frac{\mu}{30.5} \qquad \forall\, m = m_{lo}, m_{up}$$

$$\widehat{p}_m \quad \geq 0$$

The recovered probabilities for the two months m_{lo} and m_{up} are then prolonged for possible future lactations $\text{lact} = 1 \ldots N$ by a binomial expansion

$$\widehat{p}_m = \sum_{\text{lact}\,\forall\,\text{monthToLact}_{m,l}}^{N} \binom{\text{lact} \cdot (m_{up} - m_{lo})}{m - (m_{lo} \cdot \text{lact})} \cdot \widehat{p}_{m_{lo}}^{\,\text{lact}-m-(m_{lo}\cdot\text{lact})} \cdot \widehat{p}_{m_{up}}^{\,\text{lact}-m-(m_{up}\cdot\text{lact})}$$

where

$$\text{monthToLact}_{m,l} \subseteq m, l \qquad \forall\, m \geq m_{lo} \cdot l \wedge m \leq m_{up} \cdot l$$

As the comparative static version of the model FarmDyn simulates an average planning year instead of a dynamic horizon, the calving distribution needs to be adapted in order to depict an average calving coefficient over all lactations. Here, all calvings that occur in the same month of the year (e.g. January) are aggregated, and calvings after the 12th month are removed.

$$\hat{p}_m = \sum_{m1\, \forall\, m \bmod 12 \,\equiv\, m1 \bmod 12}^{M} \hat{p}_{m1} \cdot \frac{12}{l_{\text{cows}}}$$

$$\hat{p}_m = 0 \quad \forall\, m > 12$$

where

$m1 \subseteq m$

l_{cows} = Average production length of cows in months

In the model, the default proportion of males to females is assumed to be 50.8% (Foote, 1977). The calculation of the sex ratio is displayed in the following formula:

$$\frac{D_{fcalves,b,m,t}}{0.492} + 2 \cdot \delta_{acc}(n_{msex} - n_{fsex}) + 0.492 n_{fsex} - 0.508 n_{msex} = \frac{D_{mcalves,b,m,t}}{0.508} + 2 \cdot \delta_{acc}(n_{msex} - n_{fsex}) + 0.492 n_{fsex} - 0.508 n_{msex}$$

where

n_{msex} = Amount of male beef breed sexing inseminations

n_{fsex} = Amount of female sexing inseminations

δ_{acc} = Sexing accuracy

The amount of female calves (left hand side of the equation) and male calves (right hand side of the equation) are divided by their respective default sex proportion, enforcing the ratio when the equation is satisfied. With the use of sexed semen, the (default) proportion can be shifted towards more females or males respectively. However, the sexing accuracy parameter (δ_{acc}) reduces the sex bias accordingly. As a consequence of the reduced fertility of sexed sperm, the IE is assumed to increase in the model when sexed sperm is used. This both affects the number of calves being born in a year due to the (involuntarily) prolonged calving interval, as well as the input costs per successful sexing insemination.

Table S1: Assumed static input parameters for the FarmDyn model.

Parameter	Value	Source
Births per lactations	0.98	KTBL (2018)
Living calves per birth	1.04	KTBL (2018)
Calf losses (%)	5	KTBL (2018)
Average cow weight (kg)	580	KTBL (2018)
Dressing percentage old cow (%)	50	KTBL (2018)
Price old cow (€/kg meat)	2.15	KTBL (2018)
Fat content milk (%)	4.1	KTBL (2018)
Protein content milk (%)	3.4	KTBL (2018)
Milk price (€/kg ECM)	0.32	KTBL (2018)
Concentrates, 12% crude protein, (€/t)	220	KTBL (2018)
Concentrates, 18% crude protein, (€/t)	230	KTBL (2018)
Concentrates, 18% crude protein, (€/t)	270	KTBL (2018)
Soybean meal (€/t)	338	KTBL (2018)
Dystocia risk bull calves, cows (%)	7	McCullock et al. (2013)
Dystocia risk bull calves, heifers (%)	24	McCullock et al. (2013)
Dystocia risk heifer calves, cows (%)	4	McCullock et al. (2013)
Dystocia risk heifer calves, heifers (%)	14	McCullock et al. (2013)
Dystocia risk beef calves, cows (%)	8	McCullock et al. (2013)

Dystocia risk beef calves, heifers (%)	25	Assumption based on McCullock et al. (2013)
Loss from dystocia score > 3, cows (€)	73	McCullock et al. (2013)
Loss from dystocia score > 3, heifers (€)	93	McCullock et al. (2013)

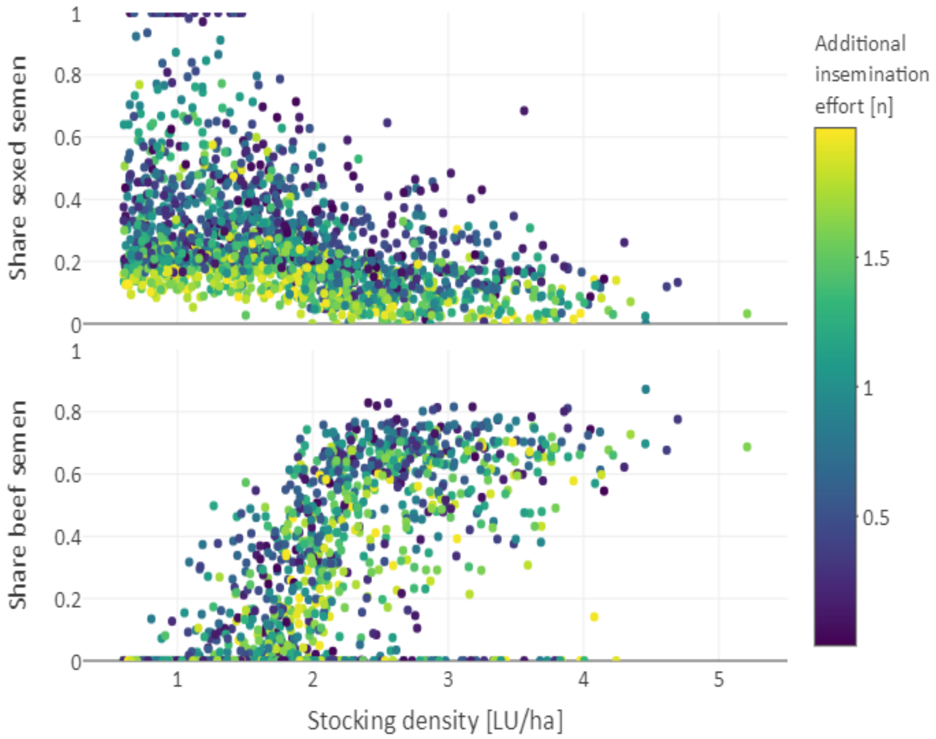


Figure S2: Profit maximizing share of female sexed dairy semen used on all inseminations (upper part), compared to the share of beef semen (sum of sexed beef semen and conventional semen) used on all inseminations (bottom part) in the North Rhine-Westphalian study population. Each dot represents a farm in the sample population. The color of the dots display the required additional insemination effort when sexed semen is used.

7.2 Appendix Chapter 3

The GAMS code of the Fruchtfolge model used in the article can be downloaded under the following permanent URL:

https://zenodo.org/badge/latestdoi/149515311

7.3 Appendix Chapter 5

Table A1: Assumed waiting periods per crop and resulting maximum crop shares in the rotation. Source: Baeumer (1990).

Crop	Waiting period (years)	Maximum crop share (%)
Field beans	4	20%
Wheat	0.5	66%
Rye	0.5	66%
Barley	0.5	66%
Maize - Corn	0	100%
Rapeseed	4	20%
Sugarbeets	4	20%
Maize - Silage	0	100%
Oats	0.5	66%

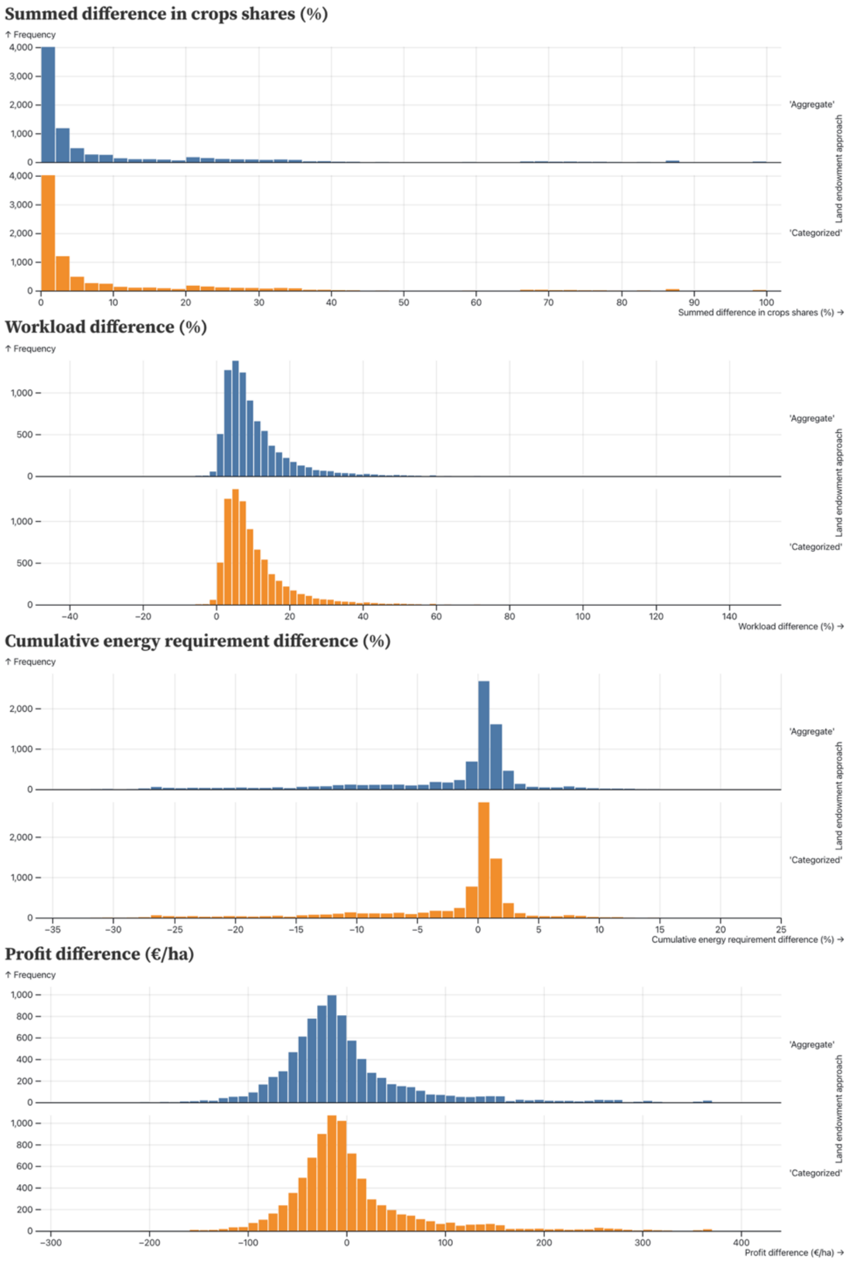


Figure A1: Histograms of differences in indicator values between the results of both the 'aggregate' (blue) and 'categorized' (orange) land endowment approach compared to the results of the 'single plot' approach.

The average, summed difference in crop shares between the 'categorized' approach and the 'single plot' approach is 11.17% (median: 2.24%, S.D.: 19.65), while the average difference in workload is 10.8% (median: 7.56%, S.D.: 11.86%). The difference in cumulative energy requirement is on average -2.03% (median: 0.4%, S.D.: 7.39%). Compared to the 'aggregate' approach, the average difference in profit is slightly lower, with an average profit difference of -1.8 €/ha, a median difference of -15.48 €/ha and a S.D. of 73.06 €/ha. A positive difference in profit indicates that simulated profits for a farm are higher in the 'aggregate' approach compared to the 'single plot' approach, and vice versa.

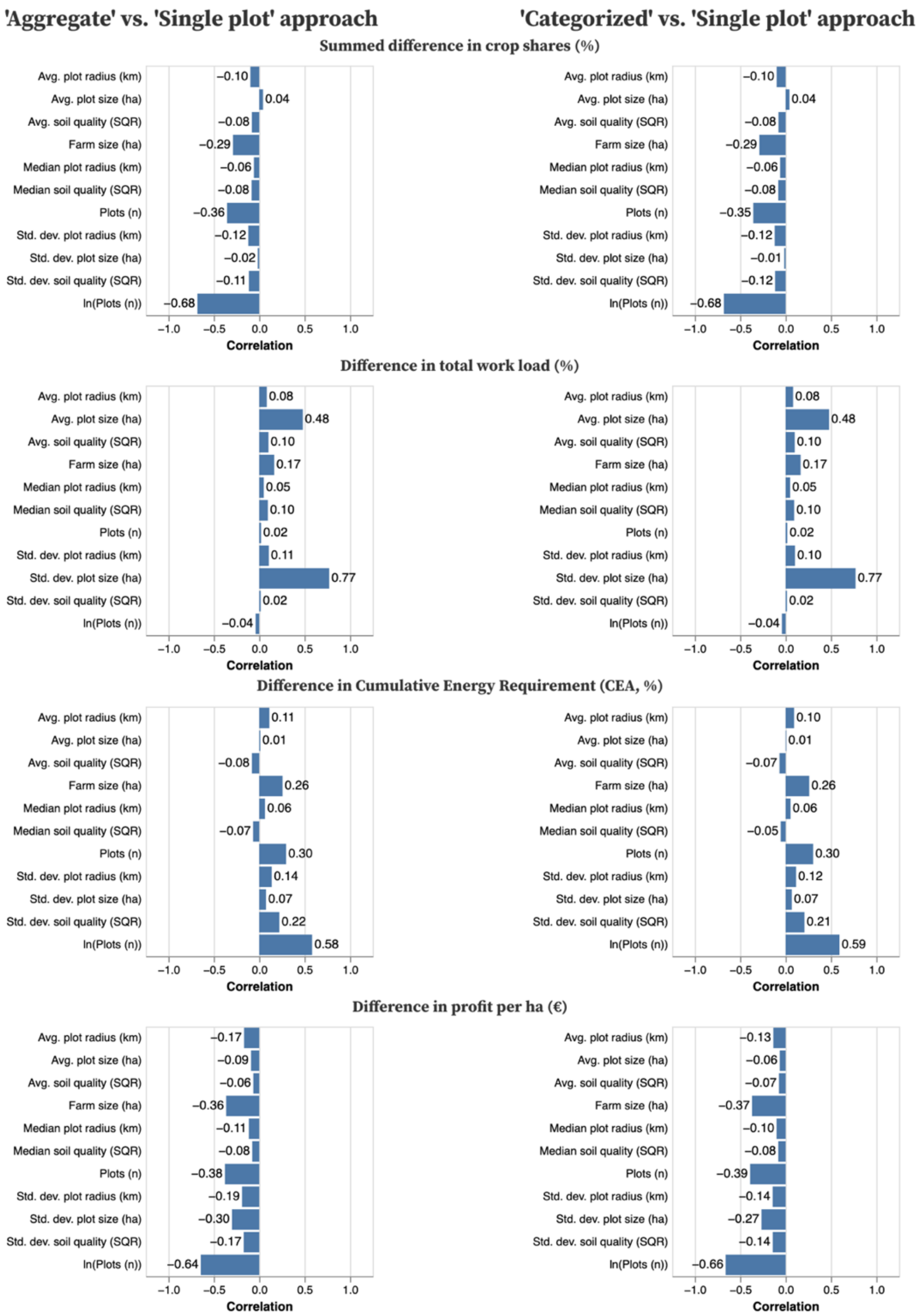


Figure A2: Pearson correlation coefficients between farm characteristics and difference in indicator values for the different land endowment approaches.

Table A2: Standardized regression results for different indicators comparing the BEFM simulations results from the ‘categorized’ and 'single plot' land endowment approach.

	Dependent variable:			
	OLS			
	‚Categorized‘ vs ‚single plot‘			
	Summed diff. in crop shares (%)	Diff. work load (%)	Diff. profit (€/ha)	Diff. CER (%)
Mean plot radius farm [km]	0.173***	-0.040***	0.115***	-0.170***
Dev. plot radius farm [km]	-0.117***	0.095***	-0.080***	0.100***
ln(Number of plots [n])	-0.752***	-0.276***	-0.647***	0.649***
Mean plot size farm [ha]	0.021*	-0.248***	0.217***	-0.011
Dev. plot size farm [ha]	0.138***	1.015***	-0.284***	-0.052***
Mean soil quality farm [SQR]	0.039***	0.026***	0.067***	-0.160***
Dev. soil quality farm [SQR]	0.028***	0.001	0.016*	0.055***
Constant	0.000	0.000	0.000	-0.000
Observations	8,509	8,509	8,509	8,509
R^2	0.502	0.684	0.477	0.399
Adjusted R^2	0.501	0.684	0.477	0.399
Residual Std. Error (df = 8501)	0.706	0.562	0.723	0.775

F Statistic (df = 7; 8501)	1,222.284***	2,629.876***	1,109.137***	807.859***

Note: *p<0.1; **p<0.05; ***p<0.01

Table A3: Regression coefficients presenting the influence of the soil quality rating (SQR) on the yield of a crop.

	Depedent variable:									
	Yield									
	OLS									
	Fieldbeans	Wheat	Rye	Barley	Maize - Corn	Rapeseed	Potatoes	Sugarbeets	Maize - Silage	Summer oats
SQR	0.264*** (0.026)	0.553*** (0.047)	0.265*** (0.056)	0.580*** (0.040)	0.486*** (0.086)	0.198*** (0.027)	4.676*** (0.392)	1.977*** (0.608)	1.743*** (0.456)	0.316*** (0.046)
Constant	18.660*** (1.580)	42.012*** (2.992)	39.906*** (3.406)	27.468*** (2.448)	64.773*** (5.407)	25.065*** (1.757)	123.584*** (24.720)	616.356*** (38.244)	335.873*** (29.087)	27.276*** (2.809)
Observations	306	306	306	306	306	306	306	306	306	306
R^2	0.247	0.309	0.068	0.404	0.095	0.150	0.319	0.034	0.046	0.132
Adjusted R^2	0.245	0.307	0.065	0.402	0.092	0.147	0.317	0.030	0.043	0.129
Residual Std. Error	1.294(*df* = 304)	1.186(*df* = 304)	1.315(*df* = 304)	1.186(*df* = 304)	1.100(*df* = 304)	1.000(*df* = 304)	1.146(*df* = 304)	1.125(*df* = 304)	1.243(*df* = 304)	1.326(*df* = 304)
F Statistic	99.874(*df* = 1;304)	136.056(*df* = 1;304)	22.135(*df* = 1;304)	205.821(*df* = 1;304)	32.057(*df* = 1;304)	53.757(*df* = 1;304)	142.501(*df* = 1;304)	10.573(*df* = 1;304)	14.612(*df* = 1;304)	46.362(*df* = 1;304)

Note: *p<0.1; **p<0.05; ***p<0.01