Institute for Food and Resource Economics

University of Bonn

Discussion Paper 2021:1

Single plots or shares of land - How modeling of crop choices in bio-economic farm models influences simulation results

Christoph Pahmeyer, Till Kuhn, Wolfgang Britz

Editor: Thomas Heckelei

Institute for Food and Resource Economics

University of Bonn

Nußallee 21

53115 Bonn, Germany

Phone: +49-228-732332

Fax: +49-228-734693

E-mail: thomas.heckelei@ilr.uni-bonn.de

Single plots or shares of land - How modeling of crop choices in bio-economic farm models influences simulation results

Christoph Pahmeyer, Till Kuhn, Wolfgang Britz

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 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

JEL classification: Q15, Q18, Q19

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 enroll 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 Programing 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.

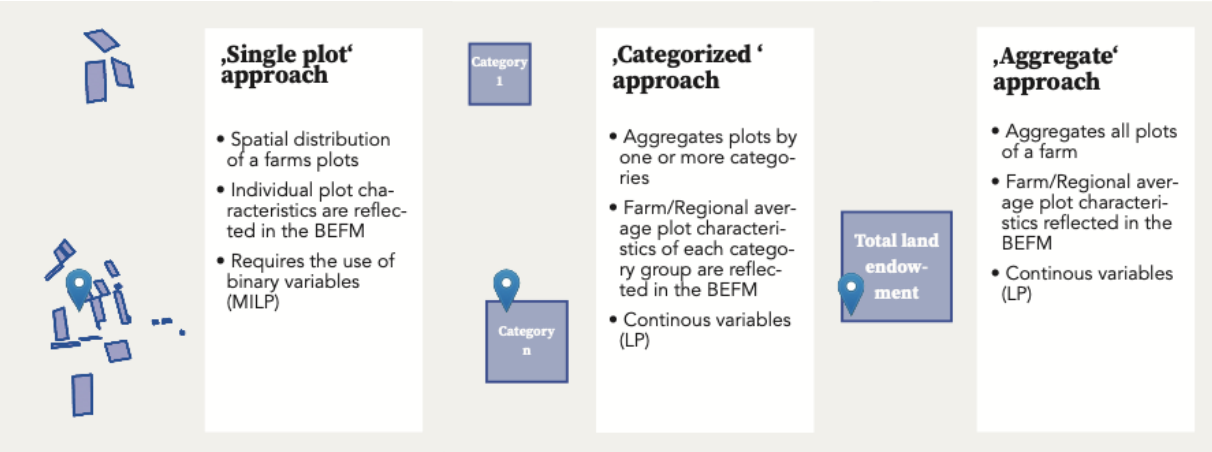
The simplified modeling of the crop choice problem based on shares of land (type) constraints likely introduces an aggregation bias. The bias is related to plot heterogeneity, i.e. the difference between mean values of plot characteristics as depicted by an aggregate constraint and the values of the 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.

2 Material and Methods

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 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 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)):

$$\max Z = \sum_{j}^{n} c_j \cdot X_j \tag{1}$$

such that

$$\sum_{j}^{n} X_j \leq b \tag{2}$$

and

$$X_j \geq 0, \text{ for all } j = 1 \text{ to } n \tag{3}$$

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. 1 – Eq. 3 may be extended as follows:

$$\max Z = \sum_{j}^{n} \sum_{s}^{o} c_{j,s} \cdot X_{j,s} \tag{4}$$

such that

$$\sum_{j}^{n} X_{j,s} \leq b_s \text{, for all } s = 1 \text{ to } o \tag{5}$$

and

$$X_{j,s} \geq 0, \text{for all } j = 1 \text{ to } n \tag{6}$$

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:

$$\max Z = \sum_{j}^{n} \sum_{k}^{m} c_{j,k} \cdot x_k \cdot V_{j,k} \tag{7}$$

such that

$$\sum_{j}^{n} V_{j,k} = 1, \text{for all } k = 1 \text{ to } m \tag{8}$$

and

$$V_{j,k} \in \{0,1\}, \text{ for all } j,k = 1 \text{ to } n,m \qquad (9)$$

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 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 1 is limited to some frequently used BEFMs in Europe, mainly drawing on the review article of Britz et al. (2021).

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[1].

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

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

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.

Table 1: Use of the different land endowment approaches in the literature, mainly based on 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) **IFM-CAP** (Louhichi et al., 2018, 2015; M'barek et al., 2017)	Regional/Sectoral policy analysis
'Categorized'	**FSSIM** (Kanellopoulos et al., 2014; Louhichi et al., 2010; van Ittersum et al., 2008) **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)	Ex-ante on-farm analysis of policy and technology adoption
'Single plot'	**MINRISK** (Radulescu and Radulescu, 2012) **FRUCHTFOLGE** (Pahmeyer et al., 2021a)	Decision support

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 'categorized' approach, crop yields are calculated for each soil type category of the farm, using the average soil quality

[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.

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 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).

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).

Figure 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.

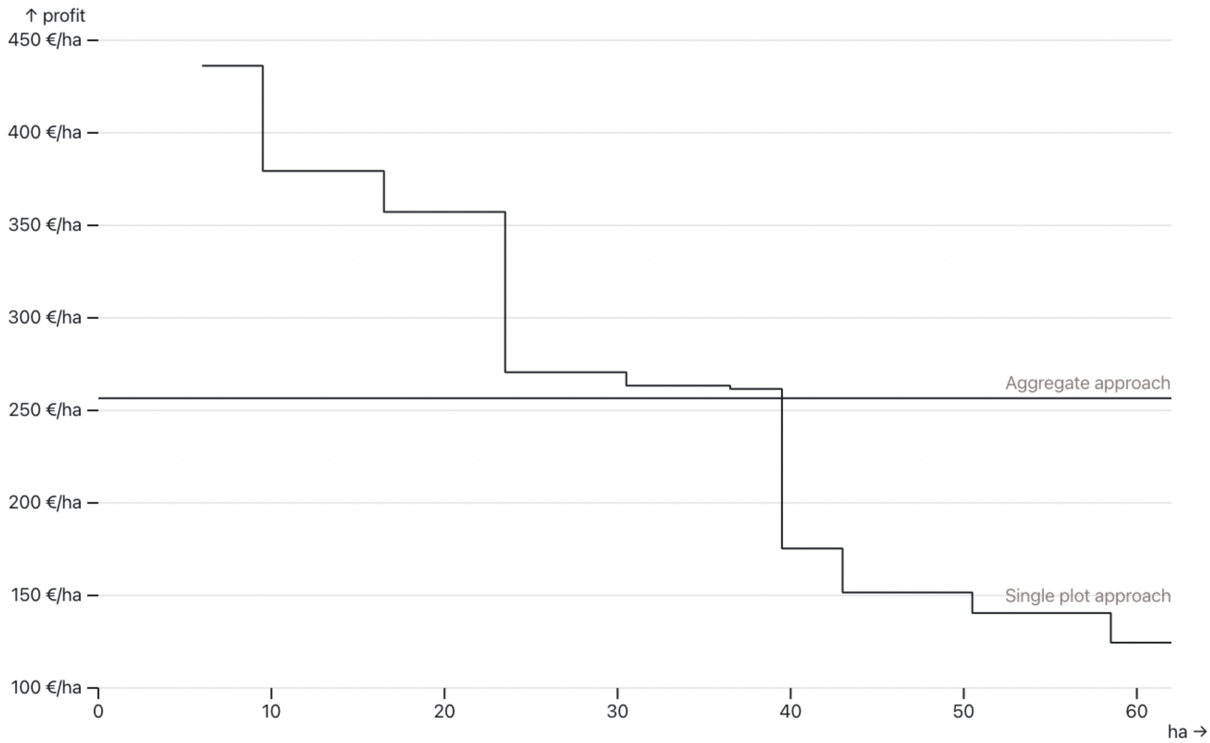


Table 2 gives an overview of the input-output coefficients for each crop, including minimum and maximum values for the case study region, reflecting varying soil qualities, farm-field-distances and plot sizes[4].

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

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

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:

$$\text{Summed difference in crop shares (\%)} = \frac{1}{2}\sum_{j}^{n}\frac{|X\_bin_j - X\_agg_j|}{b} \tag{10}$$

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

Table 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).

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

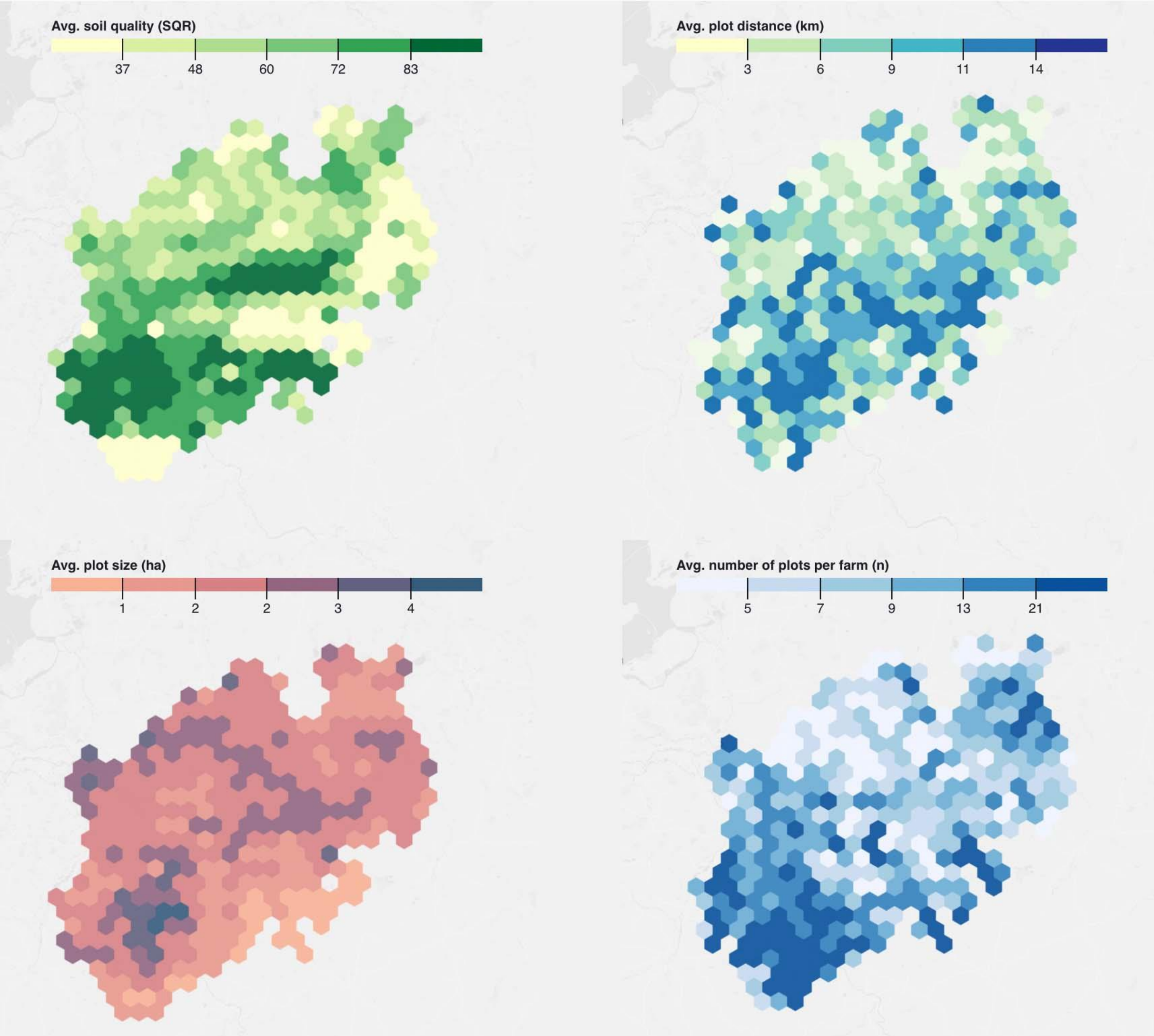
Note: For the underlying minimum and maximum values of soil quality, plot size and farm-to-field distance, see the following section (section 2.3).

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).

2.3 *Case study region*

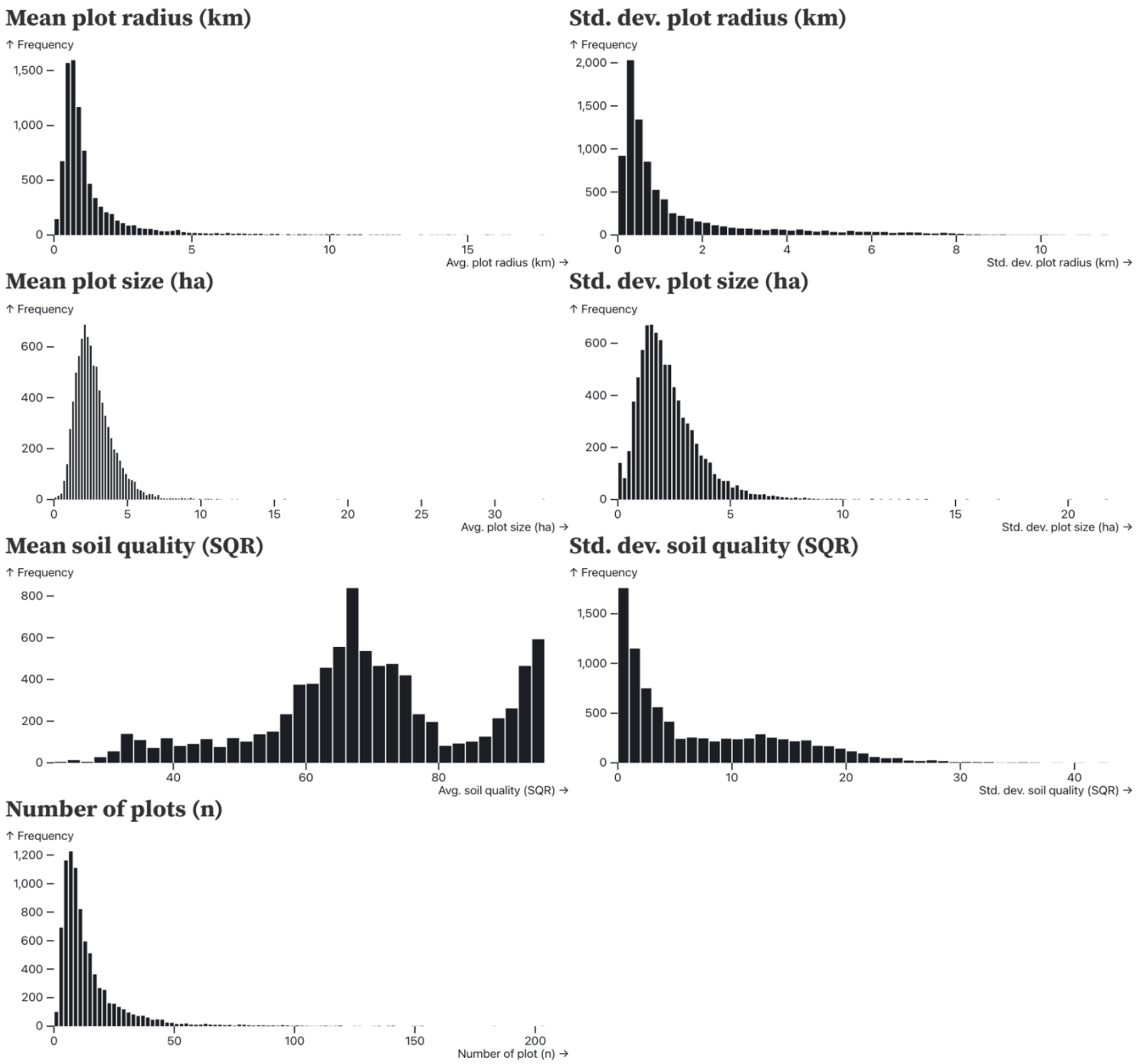
Our analysis builds on a synthetic farm population of the German federal state of North-Rhine Westphalia (Pahmeyer et al., 2021b) of which all 8,509 specialized arable farms are considered. As the classification is based on shares of revenues by farm branch in total farm revenues (Kuhn and Schäfer, 2018), 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 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 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 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.



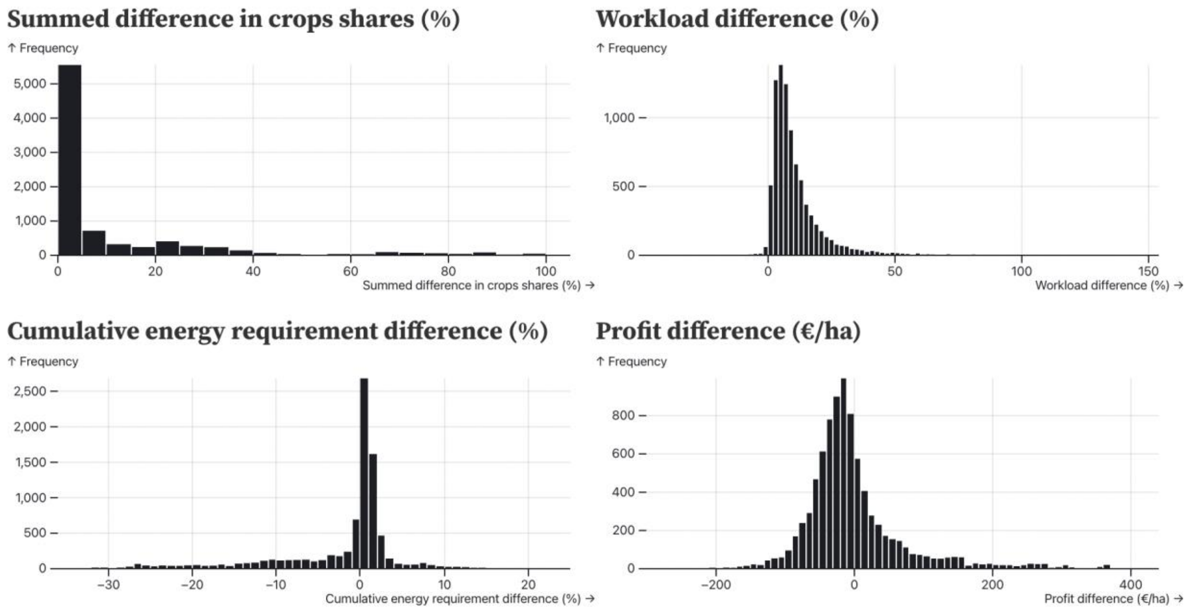
For brevity, the results section therefore focuses on the difference between the ‘aggregate’ and ‘single plot’ approach, as depicted in Figure 5. On average, the summed difference in crop shares between the two approaches is 11.15% (median: 2.23%, S.D.: 19.65%). The average difference in workload is 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 7. The histograms in Figure 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 4: Histograms of selected farm characteristic among the case-study 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.

Figure 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.



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 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 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.

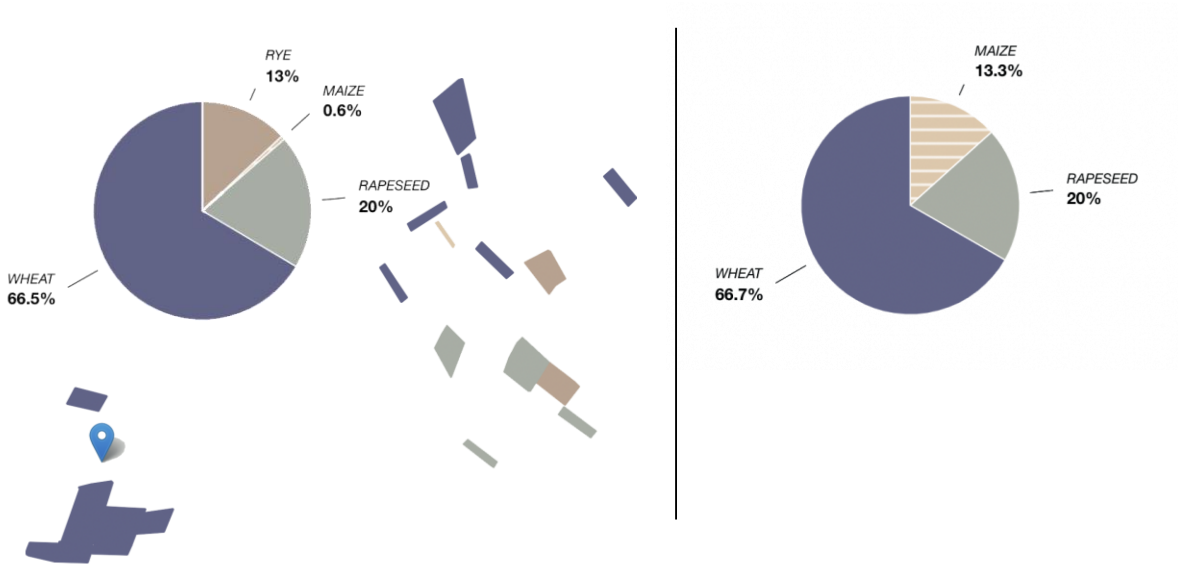


Table 3 presents the results of standardized multiple linear regression models (OLS) on differences of the four chosen indicators between the ‘aggregate’ and ‘single plot’ approach. The similar results for the ‘categorized’ approach can be found in the appendix (Table A2). For (highly) auto-correlated 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 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.

Figure 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.

Table 3: Standardized regression results (beta coefficients) for different indicators comparing the BEFM simulations results from the 'aggregate' and 'single plot' land endowment approach.

	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			

Figure 7: Total simulated cultivation area per crop in the aggregate (A, blue) and single plot (B, orange) approach.

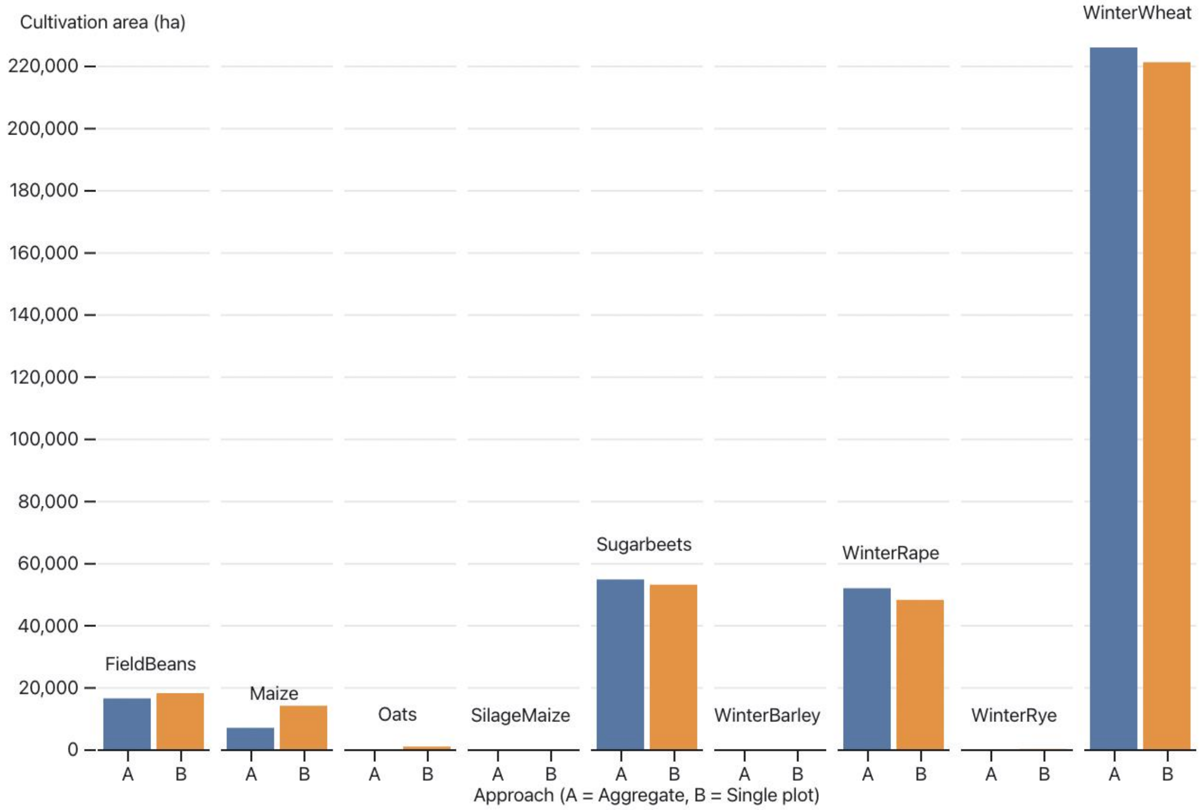


Table 4 displays the standardized regression coefficients measuring the impact of various farm characteristics on the relative difference in the total cultivation area for each crop between the 'aggregate' and 'single plot' approach. For all crops with a notable difference in the summed total cultivation 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.

Table 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.

	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 farm [SQR]	0.035***	-0.005	0.085***	0.001	-0.015*	-0.028***	-0.069***	0.004	0.120***
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

	Field beans	Wheat	Rye	Barley	Maize - Corn	Rape	Sugar beets	Maize - Silage	Oats
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:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$								

3 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.

4 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 Acknowledgements

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

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

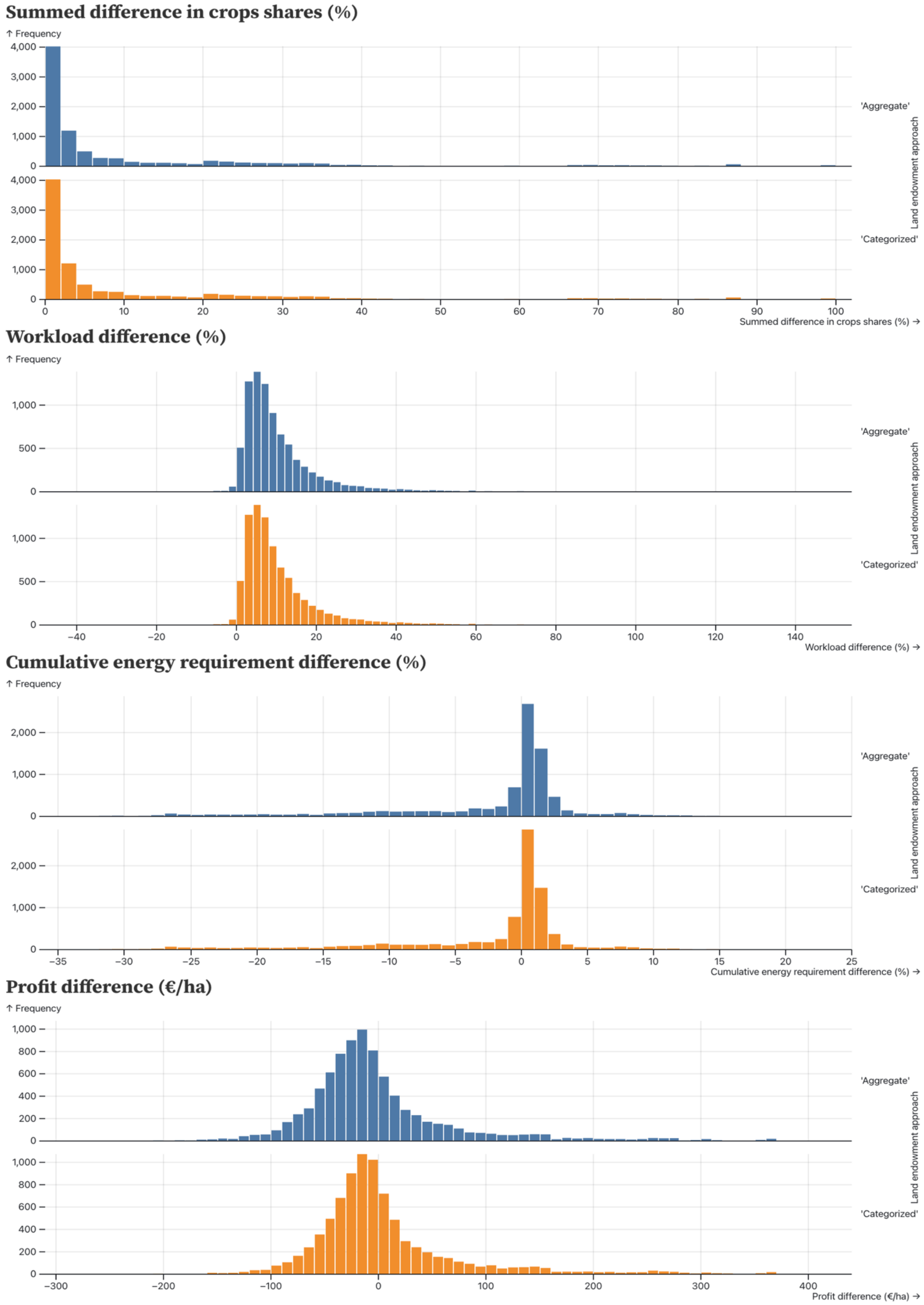
van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., van der Wal, T., Wien, J.E., Wolf, J., 2008. Integrated assessment of agricultural systems – A component-based framework for the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

Appendix A. Supplementary Data

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.

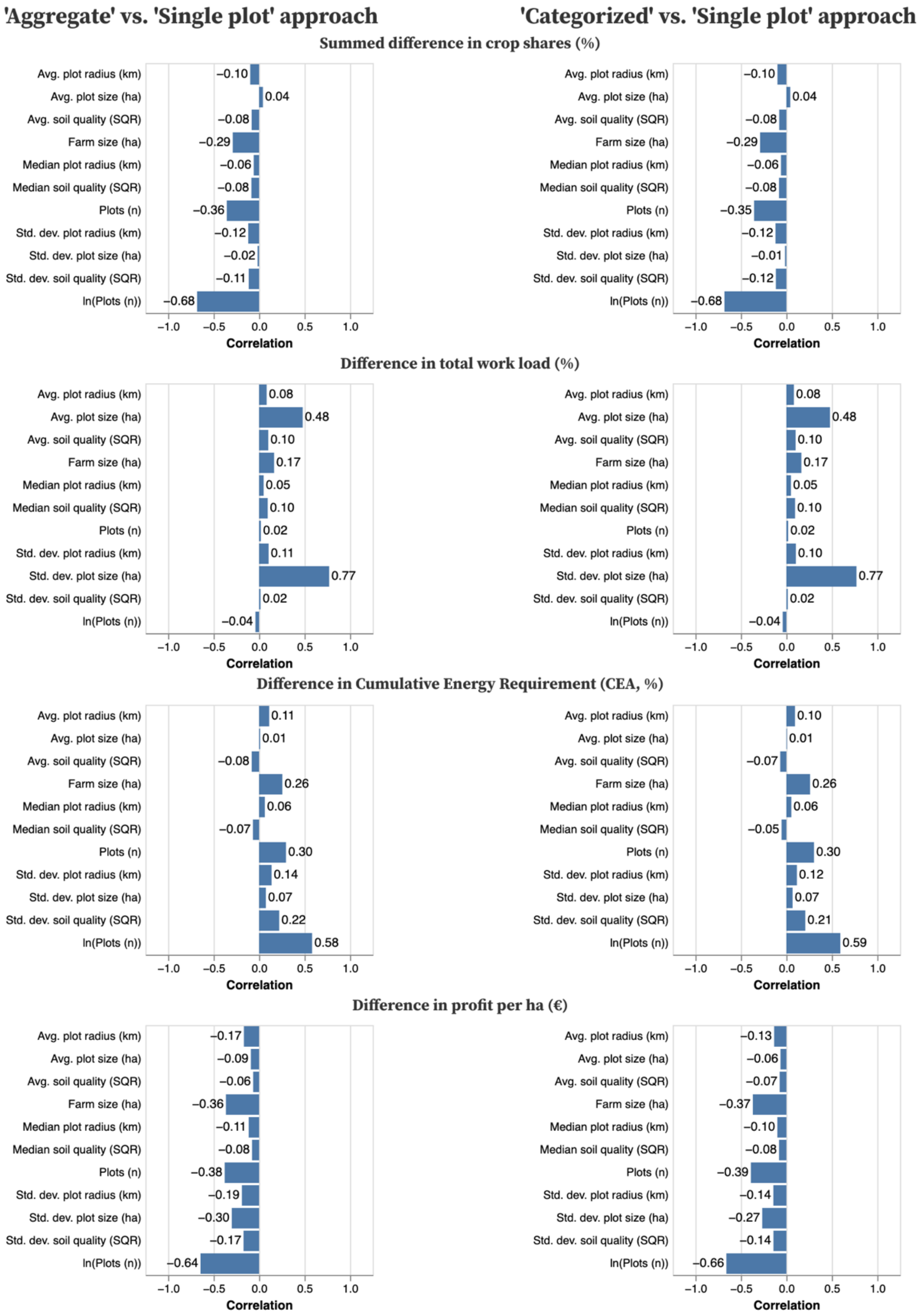


Table A2: Standardized regression results for different indicators comparing the BEFM simulations results from the ‘categorized’ and 'single plot' land endowment approach.

	Dependent variable:			
	OLS			
	‚Categorized‘ vs ‚single plot‘			
	Summed diff. in crop shares (%)	Diff. work load (%)	Diff. profit (EUR/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.

	Dependent 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