

Jänicke, Clemens et al.

Article — Published Version

Can we estimate farm size from field size? An empirical investigation of the field size to farm size relationship

Agricultural Systems

Provided in Cooperation with:

Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Halle (Saale)

Suggested Citation: Jänicke, Clemens et al. (2024) : Can we estimate farm size from field size? An empirical investigation of the field size to farm size relationship, Agricultural Systems, ISSN 1873-2267, Elsevier, Amsterdam, Vol. 220, <https://doi.org/10.1016/j.agsy.2024.104088> , <https://www.sciencedirect.com/science/article/pii/S0308521X24002385>

This Version is available at:

<https://hdl.handle.net/10419/301807>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

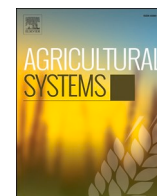
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Can we estimate farm size from field size? An empirical investigation of the field size to farm size relationship

Clemens Jänicke^{a,b,c,d,*,1}, Maximilian Wesemeyer^{b,1}, Cristina Chiarella^{e,f}, Tobia Lakes^{b,c}, Christian Levers^{g,b,h,i}, Patrick Meyfroidt^{e,j}, Daniel Müller^{a,b,c}, Marie Pratzner^{b,c}, Philippe Rufin^{e,b}

^a Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany

^b Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

^c Integrative Research Institute for Transformations of Human-Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

^d Leibniz-Centre for Agricultural Landscape Research, Eberswalder Straße 84, 15374 Müncheberg, Germany

^e Earth and Life Institute, UCLouvain, 1348 Louvain-la-Neuve, Belgium

^f International Fund for Agricultural Development - IFAD, 00142 Rome, Italy

^g Thünen Institute of Biodiversity, Johann Heinrich von Thünen Institute - Federal Research Institute for Rural Areas, Forestry, and Fisheries, Braunschweig, Germany

^h Institute for Resources, Environment and Sustainability, University of British Columbia, 2202 Main Mall, Vancouver, BC V6T 1Z4, Canada

ⁱ School of Public Policy and Global Affairs, University of British Columbia, 6476 NW Marine Drive, Vancouver, BC V6T 1Z2, Canada

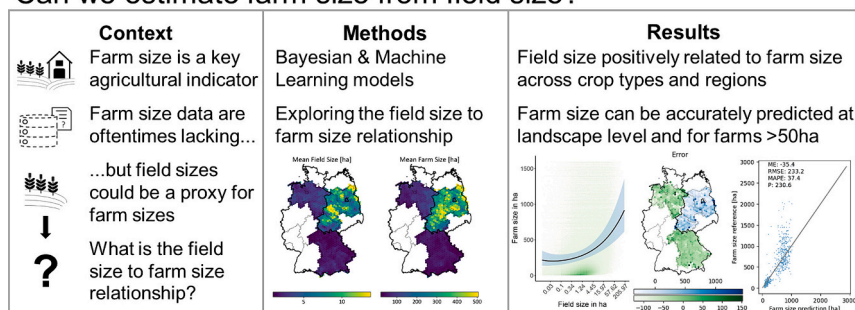
^j F.R.S. - FNRS, 1000 Brussels, Belgium

HIGHLIGHTS

- First spatially explicit analysis of the field size to farm size relationship.
- Data on 3.4 M fields and 169,000 farms representing 54% of German agricultural area.
- Positive relationship confirmed across federal states and crop types.
- Farm size reasonably well predicted in the range of 50–7,000 ha.
- Emerging opportunities for enhancing knowledge on farm size and its dynamics.

GRAPHICAL ABSTRACT

Can we estimate farm size from field size?



ARTICLE INFO

Editor: Leonard Rusinamhodzi

Keywords:

Bayes modelling
Random forest regression
Land-use

ABSTRACT

CONTEXT: Farm size is a key indicator associated with environmental, economic, and social contexts and outcomes of agriculture. Farm size data is typically obtained from agricultural censuses or household surveys, but both are usually only available in infrequent time intervals and at aggregate spatial scales. In contrast, spatially explicit and detailed data on individual fields can be accessed from cadastral information systems or agricultural subsidy applications in some regions or can be derived from Earth observation data. Empirically exploring the

* Corresponding author at: Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany.
E-mail address: jaenicke@iamo.de (C. Jänicke).

¹ Authors contributed equally to this work

<https://doi.org/10.1016/j.agsy.2024.104088>

Received 4 April 2024; Received in revised form 24 July 2024; Accepted 27 July 2024

Available online 7 August 2024

0308-521X/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Farming systems
Remote sensing
Land tenure

field-size-to-farm size relationship (FFR) is a lever to enhance our understanding of spatial patterns of farm sizes by assessing field sizes. However, our currently limited empirical knowledge does not allow for the characterization of the FFR over large spatial extents.

OBJECTIVE: We analyze the FFR using data from the Integrated Administration and Control System (IACS) for Germany. The IACS manages agricultural subsidy applications in the European Union; therefore, the data include spatial information on the extent of all fields and farms for which farmers have applied for subsidies.

METHODS: We developed a Bayesian multilevel model and a machine learning model to estimate farm size based on field size, controlling for contextual factors such as crop types, state boundaries, topography, and neighborhood effects.

RESULTS AND CONCLUSIONS: We found that farm size generally increased with field size for almost all federal states and crop type groups, but the FFR varied considerably in magnitude. Farm size predictions were accurate for medium-sized and large farms (50–7,000 ha, representing 66% of the data) with mean absolute percentage errors of 40–114%, but estimates for smaller farms had higher errors. To evaluate the relationship at the landscape level, we spatially aggregated the predictions into hexagons with a diameter of 15 km. This resulted in more accurate predictions (mean absolute percentage errors of 37%) than at the field level.

SIGNIFICANCE: Our study presents the first empirical insights into the FFR, opening future research directions towards producing spatially explicit farm size predictions at scale. Such information is key for monitoring scale transitions in agricultural systems, facilitating the design of timely and targeted interventions, and avoiding undesired outcomes of such processes.

1. Introduction

Farm size - commonly defined as the total agricultural area covered by all plots of a farm (Eurostat, 2022; Lowder et al., 2016; Samberg et al., 2016) - is a key indicator associated with environmental, economic, and social contexts and outcomes of agriculture. Farm size relates to land and labor productivity, labor intensity, land management, household economics, off-farm activities (Chiarella et al., 2023; Helfand and Taylor, 2021; Liebert et al., 2022; Meyfroidt, 2017), and biodiversity (Noack et al., 2022; Ricciardi et al., 2021). Many agricultural subsidy schemes, such as the Common Agricultural Policy (CAP) of the European Union (EU), include area-specific payments proportional to farm size, leading to competitive advantages for larger farms (Pe'er et al., 2014; Scown et al., 2020). These policies aggravate the general trend of concentration of land ownership (Williams et al., 2023), leading to increasing farm size (Debonne et al., 2022). While knowledge about farm size, its spatial patterns, and its dynamics provides valuable insights for better understanding the trade-offs between the economic, social, and environmental impacts of agriculture, the global paucity of accurate, timely, and spatially detailed farm size data represents a key empirical blind spot, which is hampering science-based design of sustainability policies in the agricultural sector.

Data on farm area size is typically obtained from agricultural censuses or household surveys. Both are challenging and expensive to collect, thus infrequently available, and rarely allow for tracking household and farm properties over time. Self-reported farm and field size in such data collections is subject to severe misreporting, such as a systematic overestimation of small and underestimation of large units (Bevis and Barrett, 2020; Carletto et al., 2013; De Groot and Traoré, 2005). Household surveys tend to underrepresent larger farms and non-family farms and often fail to capture information about farms from owners who reside elsewhere (Jayne et al., 2022). Besides, few censuses or household surveys collect spatially explicit information, such as GPS locations of fields or field boundaries. Moreover, for privacy reasons, most datasets report farm size in aggregated form at the national level or for subnational administrative units. Such spatially aggregated data does not allow to assess patterns, dynamics, causes, and impacts of farm size, and analyses at aggregate levels are prone to the ecological fallacy (Meyfroidt, 2017).

However, many land use and land management variables correlate with the size of farms. For instance, the presence of a particular crop type or land management practice may help distinguish different farming systems that correspond to different farm size categories (e.g., small-scale semi-subsistence farming versus commercial farming by larger farms) (Bey et al., 2020). Similarly, field sizes (or categories

thereof) may relate to categories of actors defined by farm size (Branthomme et al., 2023; Dang et al., 2019). In many high-income countries, cadastral data or data for agricultural subsidy applications are systematically collected, for example, in the EU via the Integrated Administration and Control System (IACS) (Leonhardt et al., 2023). The IACS administers and controls farmers' applications for agricultural subsidies, whereby they must specify the spatial extent and location of each plot, the crop planted on each field, as well as basic farm-level variables, such as organic or conventional farming. The IACS data, therefore, contain field-level information on field size and land use. A unique identifier of the subsidy applicant allows assigning fields to farms and thus, the calculation of farm sizes. The paying agencies of each EU country (or each federal state in some countries) manage the subsidy applications. The IACS data, and especially the farm identifiers, are generally not publicly available but some countries release the IACS data without the farm identifier. Additionally, they may be requested for scientific purposes. Due to the decentralized data collection system, the IACS data requires harmonization to be consistent across jurisdictions. Despite these drawbacks, IACS data represent a unique data source for assessing dynamics, patterns, and impacts of field and farm sizes. For regions where such data is not available, recent developments in Earth observation (EO) and computer vision also present emerging opportunities to produce spatially detailed and timely estimates of field size (Waldner and Diakogiannis, 2020) and other spatially explicit information on land use (Kuemmerle et al., 2013; Verburg et al., 2011).

Leveraging field size estimates for farm size estimation requires understanding the field size to farm size relationship (hereafter: FFR). The current evidence materializes as two general observations stating (1) a generally positive relationship between farm size and field size, where farm size naturally puts an upper bound on the size of fields therein, and (2) a notable regional variability and scale dependency in the relationship between both measures (Levin, 2006). However, the strength and shape of the relationship between field and farm size and the predictor variables describing the contextual settings influencing this relationship are unknown. Farm and field size vary across space and time, influenced by many factors. Farm-level decisions on allocating agricultural land into fields are path-dependent and include choices and constraints regarding crop types, farm-level crop diversity, available labor, inputs (e.g., fertilizer and pesticides), mechanization, livestock integration, or irrigation. Landscape-level biophysical factors, such as soil types, slope, water availability, climate, population density, or the quality and availability of agricultural land shape within-farm land allocation. At the administrative level, agricultural policies, land reforms, land prices, accessibility, inheritance patterns (and hence culture and demography), ownership structure, and land-use history may influence the division of

agricultural landscapes into farms and fields of various sizes (Dannen-berg and Kuemmerle, 2010; Sikor et al., 2009). Consequently, the FFR can be expected to vary in space and time but remains unobserved to date.

The paucity of farm size data and the increasing availability of IACS data or EO-based field delineations call for assessing the opportunities for using field size as a proxy for farm size. The overarching aim of this study is to analyze the FFR and its variation in a setting with very good data availability on field- and farm sizes and different predictor variables describing the contextual settings. We define farm size as the total area covered by all farm plots. We selected Germany as our study region due to the availability of data from the Integrated Administration and Control System (IACS), which contains all farms and fields for five federal states in Germany. We have three specific objectives. First, we aim to understand the statistical relationship between farm size, field size, and other predictor variables describing the contextual settings using a Bayesian multilevel model. Second, we estimated farm size using the Bayesian multilevel model and a machine learning regression model to assess the influence of methodological choices on the quality of farm size estimates. Third, we produced an in-depth error diagnosis to assess

which farm sizes we can (or cannot) produce estimates with acceptable uncertainty.

2. Data & methods

The pre-processing and analysis followed the workflow represented in Fig. 1. In total, we conducted three analyses: First, we developed a Bayesian multilevel model to assess the statistical relationship between farm size, field size, and additional predictors in the study region. Second, we trained a random forest regression model (RFR) to estimate farm size with the same predictor variables to compare the performance of the RFR with the Bayesian model. Third, we assessed the accuracy and precision of farm size predictions obtained from both models. This involved quantifying the uncertainties associated with the predictions and the heteroscedasticity across farm sizes. The entire analysis is available online at: https://github.com/clejae/field_farm_size_relationship.

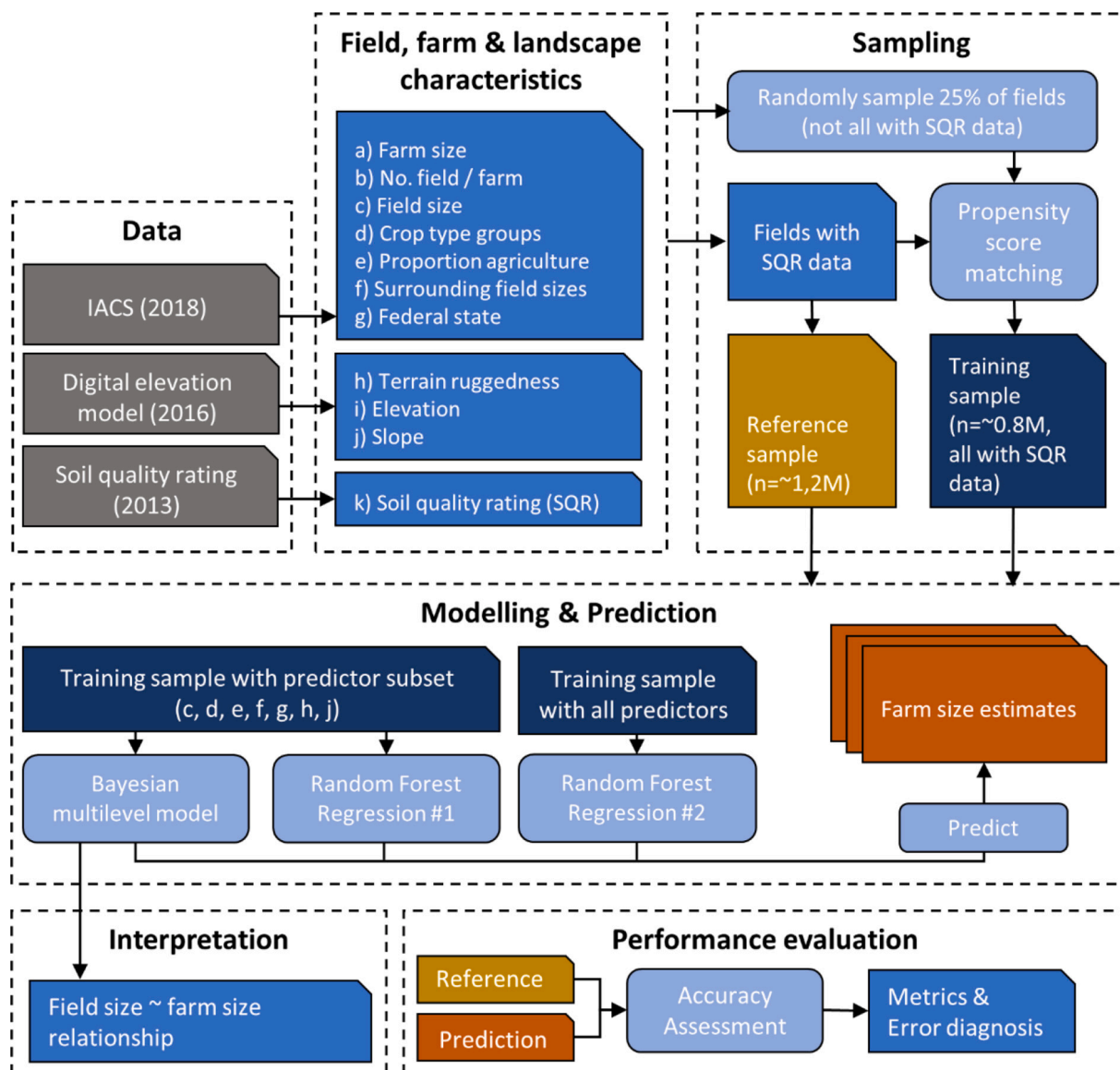


Fig. 1. Overview of the processing workflow. All boxes with rounded corners represent processing steps, e.g. sampling, matching or prediction, while all boxes with sharp corners represent datasets, e.g. training or reference samples or estimates.

2.1. Field & farm data

Our study area comprises five federal states in Germany, namely Bavaria, Brandenburg, Lower Saxony, Saxony-Anhalt, and Thuringia. The study area covers more than 90,000 km² of agricultural land, or 54% of the total agricultural area of Germany (Fig. 2). Bavaria and Lower Saxony were part of the former Federal Republic of Germany (henceforth West Germany); Brandenburg, Saxony-Anhalt, and Thuringia were part of the former German Democratic Republic (henceforth East Germany).

We used data from the Integrated Administration and Control System (IACS) reported in EU subsidy applications for 2018. The data contains 3.4 million field delineations of all cropland and grassland fields with a farm identifier (Table 1). Farm and field sizes and the number of farms and fields vary considerably across the federal states, especially between East and West Germany (Fig. 3). The average field and farm sizes in the East German states are 8.0 ha and 225.0 ha, respectively, compared to 1.6 ha and 29.9 ha in the West German states. Our data comprises 153,041 farms in West Germany and 15,967 farms in East Germany.

We removed small landscape elements and objects that were delimiting features, such as hedges or tree rows, as these landscape elements are unlikely to be recognized as a field in a remote sensing study aimed at deriving field boundaries from satellite imagery. Approximately 46% of the total number of fields are planted with crops and 43% are grasslands; the remaining fields are other types of land cover or land use, such as permanent crops, flower strips, or herbal fields. Crop fields are on average more than twice as large as grassland fields (4.0 ha vs. 1.8 ha), and account for a larger total area (61,950 km² cropland and 25,933 km² grassland). Each field has a farm identifier that allows computing the size of all farms ($N = 169,203$).

For each field and year, the data contain information on more than 400 crop types planted. We aggregated crop types into crop-type groups based on similar functional types and similar average field sizes (Table 2).

Farms can operate on fields that are located outside of the federal state where the farm is registered. This has two effects: First, the number of farms is inflated as the farm identifiers are not uniform across all federal states. In Brandenburg, Saxony-Anhalt, and Thuringia the farm identifiers are the same across state borders, so that farms can be traced

across state borders. In Bavaria and Lower-Saxony the farm identifiers follow different principles. Thus, it is possible that farms from Bavaria and Lower-Saxony that also have fields in one of the other states, or vice versa, are artificially split into several farms. The second effect is that farm sizes are underestimated in some cases. This may happen because some fields, which belong to a farm within our study region, are in federal states outside of our study region. Another case is that fields within our study region could belong to farms located outside of the study region. To assess the magnitude of this effect, we calculated the number of farms which operate across federal states for Brandenburg, Saxony-Anhalt, and Thuringia, while this was not possible for the West German states. We found that 12% of farms operated in more than one state; only accounting for the primary location of farm property would lead to an average farm size underestimation of 13.8% ($\pm 13.3\%$ standard deviation). These farms were mostly located in proximity to administrative boundaries.

We hypothesized that the strength of the association between farm size and field size increases when a farm has fewer fields, up to the extreme case with one field per farm where farm size equals field size. We conducted simulation experiments (see Appendix 1) with randomly generated field and farm data to confirm this hypothesis. We found that the positive association is indeed stronger for farms that have fewer fields and this association approaches zero the more fields a farm has. To avoid potential biases from this effect, we modeled farm size independently from the size of the field under investigation by subtracting the area of each field from associated farm size. Consequently, our outcome variable represents the remaining farm size excluding the area of the field of interest. Importantly, while this approach was primarily implemented to avoid biases in the model introduced by the (unknown) number of fields for each farm, the total farm size can be obtained by simply adding the prediction of remaining field size and the size of the field under study. In the cases when a farm had only one field (7,249 farms in our dataset), the model predicted a value of remaining farm size close to zero, in line with expectations.

2.2. Predictors

We compiled a set of twelve predictors to model the FFR. The predictors were selected based on scientific literature, data exploration, or hypothesized relationships with farm size (Table 3). Partly, the selection was constrained by data availability. For example, data on land prices, land management (e.g., livestock densities, irrigation, fertilizer input), or designated markets (crop production for food, fodder, or bioenergy) were hypothesized to have an influence on the FFR but were not available in the sufficient spatial detail and timeliness for our analyses. Further, the number of livestock, organic agriculture, or other farm-related management indicators arguably can influence field size distributions, but farm-level data covering all states were unavailable to us and we hence had to exclude these variables from the analyses.

For data preparation, we first log-transformed the predictor variables field size and average surrounding field size to reduce skewness and improve symmetry. To facilitate model convergence and setting meaningful priors, we scaled and centered all predictors including the log-transformed ones to zero mean and unit standard deviation. We detected correlation between terrain ruggedness and surrounding field sizes (-0.5), which changed the direction of the effect of terrain ruggedness (Appendix 4). We therefore assumed a confounding effect of this variable pair and excluded terrain ruggedness as we considered surrounding field size more important for estimating farm size.

We used a 25% sample from the entire dataset for training our models. As the soil quality rating dataset had spatial gaps, a simple random sample would have included fields with no information on soil quality. We suspected the missing information to not be randomly distributed. To obtain a random sample without data gaps in soil quality we first drew a random sample that contained fields with no information on soil quality. Then we created a new dataset with only fields that

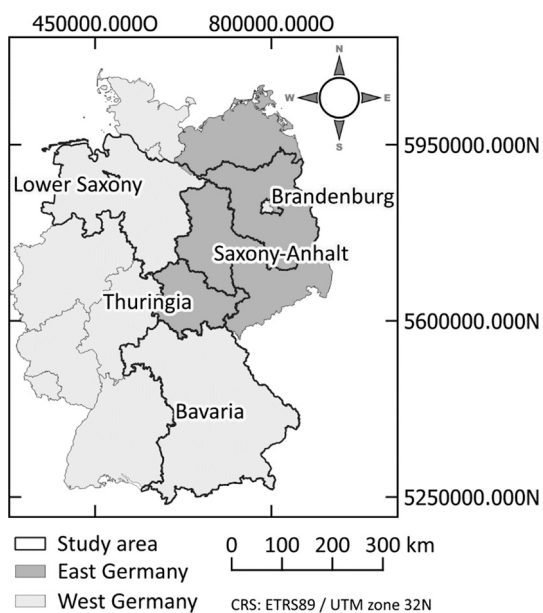


Fig. 2. Overview map of the study area. East Germany refers to the federal states of the former German Democratic Republic. West Germany refers to the states of the former Federal Republic of Germany.

Table 1

Overview of the number of fields and farms in IACS data per state. The farm numbers refer to the farms that operate in the respective state and not to the number of farms that have their farmstead in the state.

State	Region	Number of fields	Number of farms	Average farm size [ha]	Average field size [ha]	Total agricultural area [km ²]
Bavaria	West Germany	1,950,442	105,844	30	1.6	31,653
Lower-Saxony	West Germany	905,247	47,197	55	2.9	25,840
Brandenburg	East Germany	165,630	6,032	225	8.0	13,300
Saxony-Anhalt	East Germany	138,990	4,897	245	8.5	11,830
Thuringia	East Germany	141,062	5,038	151	5.7	80,705
All West German states		2,855,689	153,041	29.9	1.6	57,493
All East German states		445,682	15,967	225.0	8.0	33,200
Total		3,382,990	169,203	54	2.6	90,836

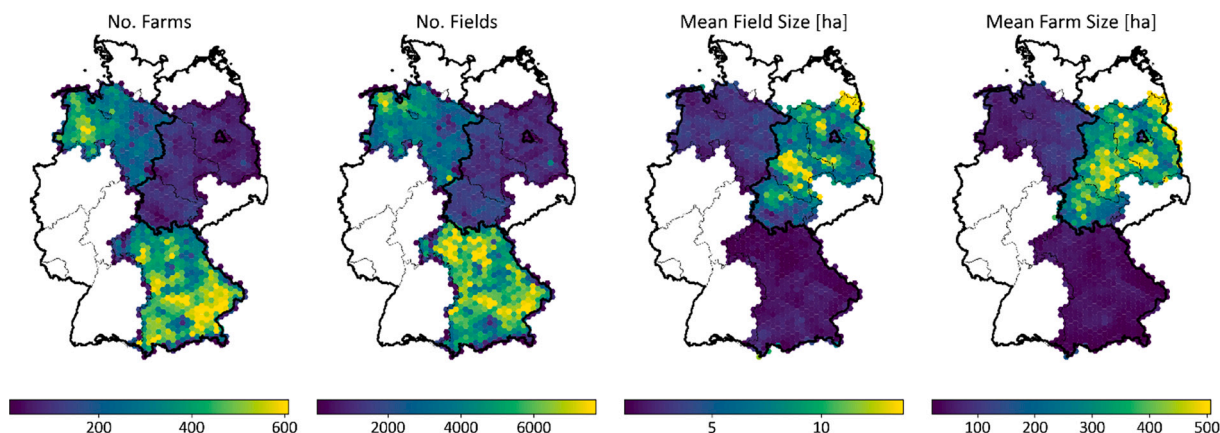


Fig. 3. Number of farms, number of fields, mean field size (ha), and mean farm size (ha) for Bavaria, Brandenburg, Lower Saxony, Saxony-Anhalt, and Thuringia. Data are aggregated to hexagons with a vertical and horizontal diameter of 15 km (~195 km²). Grey lines indicate federal states of Germany and thick black line indicates the former border between West and East Germany as well as the national border of Germany.

contained soil quality information and then used propensity score matching (PSM) to find matches for the random sample. This allowed us to construct a sample that was not strictly drawn randomly, but that was very similar to the random sample, and with consistent soil quality information for each observation. For more details on the PSM, see Appendix 5.

2.3. Bayesian model description

We aimed at investigating the effect of field size and the selected predictor variables on the magnitude and variance of farm size. To represent the dependency of this relationship on other predictors, such as administration and management, we allowed for additional variance at the level of federal state and crop type. Therefore, we used a Bayesian multilevel modelling approach, which offers two main advantages. Firstly, it allowed us to estimate parameter uncertainty by estimating probability density functions instead of point estimates for parameter

coefficients. Consequently, we could assess the likelihood and related uncertainty of predicted farm sizes for every specific field size. Secondly, it allowed us to implement a multilevel model structure, which offers flexibility for modelling statistical phenomena that occur on different levels, such as the farm level and higher administrative levels. Each level has its own set of parameters, which are influenced by the estimates at another level. This permits to capture both aggregate trends and local variations in the relationships between predictors and the response.

The model's overarching goal is to understand the relationship between field size and farm size while accounting for crop-specific variability and state-specific differences, as well as the field size and proportion of agricultural area. Therefore, we modeled farm size as a function of a lognormal distribution, with its mean value and the standard deviation being a function of our predictor variables. To capture the variability across different crops and states, we included group-level intercepts and slopes, as well as group-level intercepts for sigma. These group levels are based on the combination of federal state and crop type.

Table 2

Crop type groups with their acronyms and a description of the included crops.

Crop type group	Acronym	Description	Average field size in study area [ha]	Average field sizes in West German states [ha]	Average field sizes in East German states [ha]
Maize	MA	All types of maize, e.g., silage maize, grain corn, or maize for biogas product etc.	3.53	2.77	13.15
Cereals	CE	Summer and winter barley, rye, wheat, and triticale.	4.14	2.71	12.78
Oilseed rape	OR	Summer and winter rape.	6.48	3.10	16.20
Beets	SU	Sugar beet, fodder beet, and other beets.	5.48	4.27	17.82
Potato	PO	Table potatoes, seed potatoes, and starch potatoes.	3.69	3.34	9.33
Legumes	LE	Extensive and intensive legumes and lupins, peas, vetches, lucernes, clover, etc.	2.48	1.69	7.10
Grasses	GR	Arable grasses and permanent grassland.	1.78	1.50	3.65
Other	OT	Other crops (e.g., herbals or flowers), permanent crops, vegetables etc.	0.96	0.63	2.22

Table 3

Predictor variables for farm size with expected effects on farm size, underlying hypotheses, description of the variables, and data sources. A “+” stands for a positive expected effect of the predictor on the company size, a “-” for a negative expected effect and a “±” means that we expect an effect that can go in either direction, depending on the value of the predictor.

Predictor	Effect on farm size	Reasoning & description	Data source
Field size	+	We assume a positive correlation between field size and (remaining) farm size, beyond the mere mathematical fact that for any given field size, the farm to which it belongs cannot be smaller than the size of that field (which indicates that larger fields result in larger farms).	IACS 2018
Crop type groups (<i>new_IDKTYP</i>)	± (group specific)	Field sizes differ between farms with different specializations, hinting towards different field sizes between crop types (Roschewitz et al., 2005). We assume that different crop types are typically characterized by distinct forms of management, levels of mechanization, and economies of scale, leading to different field size across crops. We indeed observed a strong variation in field sizes among the different crop type groups (Table 2) and in the relationship between field sizes and farm sizes (Appendix 2).	IACS 2018
Federal state (<i>federal_st</i>)	± (group specific)	Field and farm sizes in Germany are determined by political legacies. While family farms were strongly favored in West Germany, large agricultural cooperatives and state farms emerged in the East Germany (Wolz et al., 2010). These differences are still reflected in the average farm sizes today (Noack et al., 2022). Additionally, other state-level policies could be captured by this variable. We represent this in our model using a categorical variable indicating in which federal state the field is located.	IACS 2018
Surrounding field sizes (<i>surrf_mean</i> , <i>surrf_std</i> , <i>surrf_min</i> , <i>surrf_max</i>)	+ (<i>surrf_mean</i> , <i>surrf_min</i> , <i>surrf_max</i>) ± (<i>surrf_std</i>)	There are large variations in field sizes within the federal states. Our explorative analysis revealed that large farms, such as those found in East Germany, also cultivate small fields close to their large fields. For these small fields, we therefore deduce that the size of the neighboring fields can serve as an indicator of farm size. We considered the mean, standard deviation, minimum and maximum of the field sizes within 1000 m distance of each field.	IACS 2018
Local proportion of agricultural area (<i>proportAgr</i>)	+	Non-agricultural land uses, such as forestry and urban areas, compete with and limit the available agricultural land in a region. We assume that increasing land-use competition and limited availability of agricultural land constrains the presence of large farms and fields (although contextually this may vary, as described in Appendix 3). We considered the proportion of agricultural area within a 1000 m radius of each field.	IACS 2018
Terrain ruggedness (<i>TRI_avg0</i> , <i>TRI_avg1000</i>)	-	We assume that rugged terrain (as a measure of elevation differences in a region) limits the possibility of creating larger fields and farms. We considered the mean terrain ruggedness for each field and within a 1000 m radius of each field.	(European Environment Agency (EEA), 2016)
Elevation (<i>ElevationA</i>)	-	We assume that elevation influences farm and field sizes. We assume smaller sizes with increasing elevation due to increasingly suboptimal agroclimatic conditions. We considered the mean elevation for each field.	(European Environment Agency (EEA), 2016)
Soil quality rating (<i>SQRAvg</i>)	±	Soil quality has a positive impact on farmland prices (Meissner and Musshoff, 2022), and the best soils tend to be acquired by farms with high capital endowments. To take advantage of economies of scale, these farms establish large fields on these soils. Additionally, soil quality can have an indirect effect on farm size, as better soils produce higher yields, which leads to higher profits and eventually larger farms. However, low soil quality can also have a positive effect on the size of fields, as larger areas of low quality are required to achieve similar overall production levels (although contextually this may vary, as described in Appendix 3).	(BGR, 2013)

We conducted a model selection using leave-one-out cross-validation and reported our results based on the best selected model. This model included the variables field size, surrounding field size, crop type group, federal state, and proportion of agriculture.

We used a lognormal likelihood function to model farm sizes (FS) with a location parameter μ and a scale parameter σ (Eq. 1).

$$FS_i \sim \text{Lognormal}(\mu_i, \sigma_i) \quad (1)$$

The location parameter μ controls the location of the peak of the distribution whereas a higher value for the scale parameter σ translates to a larger spread or width of the distribution. Note that increasing location parameter values lead to a wider spread on the outcome scale, too. The priors for the parameters were set on the log scale while the predictions were on the original scale (farm size in ha). The location and the scale parameter were modeled in separate linear models described below.

The location parameter μ was modeled with a multilevel linear model of the predictors (Eq. 2).

$$\begin{aligned} \mu_i = & \bar{\alpha} + \alpha_{\text{STATE_CropT}[i]} + \beta_{\text{STATE_CropT}[i]} \log\text{FieldSize}_i + \beta_1 \log\text{FieldSize}_i \\ & + \beta_2 \text{sur_fieldsize}_i + \beta_3 \text{propAg1000}_i \\ & + \beta_4 \text{propAg1000}_i * \text{sur_fieldsize}_i + \beta_5 \text{SQRAvg}_i \end{aligned} \quad (2)$$

The model included a global intercept (Eq. 3) and population level parameters (Eq. 4). We set an informed prior for the global intercept parameter at 4.5 (Eq. 3), which roughly represents the average log farm size in the 5 federal states.

$$\bar{\alpha} \sim \text{Normal}(4.5, 1) \quad (3)$$

$$\beta_1, \dots, \beta_5 \sim \text{Normal}(0, 0.5) \quad (4)$$

Additionally, we allowed the intercepts and slopes of field sizes to vary by federal state and crop type group. For the sigma parameter reflecting the varying intercepts (Eq. 5), we set an informed prior at *Halfnormal*(1, 0.5) (Eq. 6) to reflect the considerably varying farm size between the federal states and crop type groups. The smallest farms are in Bavaria with a log farm size of 3.4 and the largest in Thuringia with 5.6. The prior reflects these differences because it allows deviations from the global intercept to cover both Thuringia and Bavaria. Based on data exploration (Fig. 3), we assumed the variation of farm size between states to be greater than between crops. The varying intercepts can be interpreted as the deviation from the global intercept (Eq. 3).

$$\alpha_{\text{STATE_CropT}[i]} \sim \text{Normal}(0, \sigma_{\alpha\text{STATE_CropT}[i]}) \quad (5)$$

$$\sigma_{\alpha\text{STATE_CropT}[i]} \sim \text{Halfnormal}(1, 0.5) \quad (6)$$

We introduced varying slopes (Eq. 7) to allow the effect of field size to vary by crop type and state. For instance, we suspected grassland fields to have a different relationship with farm size compared to, e.g., potato fields, and that this differential effect also varies across states. We thereby explicitly include the fact that some states have smaller farms, and that some crops are cultivated on larger fields than others. The interaction between the two categorical variables for the federal state and the crop type groups yielded 40 combinations (“STATE_CropT”), and we modeled intercepts and slopes for each of these combinations. The scale parameter of the varying intercepts was set with a prior of *Halfnormal*(0.5, 0.5) (Eq. 8). Again, the varying intercepts represent the deviation from the population level (Eq. 4).

$$\beta_{\text{STATE_CropT}[i]} \sim \text{Normal}\left(0, \sigma_{\beta_{\text{STATE_CropT}[i]}}\right) \quad (7)$$

$$\sigma_{\beta_{\text{STATE_CropT}[i]}} \sim \text{Halfnormal}(0.50.5) \quad (8)$$

The scale parameter σ was modeled with a varying intercept linear model, i.e., a multilevel model (Eq. 9). We used a log link function to ensure that sigma is positive. The varying intercept linear model allowed us to understand the uncertainties of the farm size predictions. The scale parameter σ is a function of field sizes, of a global intercept (Eq. 10), and of the varying intercepts that vary by federal state and crop type groups (Eqs. 11 & 12). To deal with heteroscedasticity in the data with increasing field sizes, we added β_{σ} (Eq. 13), which allows sigma to change with increasing field sizes.

$$\text{Log}(\sigma_i) = \bar{\alpha}_{\sigma} + \alpha_{\text{STATE_CropT}[i]} + \beta_{\sigma} \log \text{FieldSize}_i \quad (9)$$

$$\bar{\alpha}_{\sigma} \sim \text{Normal}(1, 0.5) \quad (10)$$

$$\alpha_{\sigma_{\text{STATE_CropT}[i]}} \sim \text{Normal}\left(0, \sigma_{\alpha_{\sigma_{\text{STATE_CropT}[i]}}}\right) \quad (11)$$

$$\sigma_{\alpha_{\sigma_{\text{STATE_CropT}[i]}}} \sim \text{Halfnormal}(0.5, 0.5) \quad (12)$$

$$\beta_{\sigma} \sim \text{Normal}(-0.1, 0.2) \quad (13)$$

2.4. Implementation

To obtain the parameter estimates, we sampled 4,000 realizations of the posterior distribution using Markov Chain Monte Carlo with four sampling chains running for 2,000 iterations and a warm-up period of 1,000 iterations each. Convergence was verified using the R-hat statistic and examination of trace plots. We evaluated model fit based on prior and posterior predictive checks, that is, predicting new hypothetical data sampled from the posterior predictive distribution and comparing it to a random draw of observed data (Appendix 6 & Appendix 7). Further, we determined out-of-sample predictive accuracy using the leave-one-out expected log predictive density method with the R package loo (Vehtari et al., 2021). We performed all modelling through the brms package in R (Bürkner, 2018) as an interface to the Bayesian inference engine Stan (Carpenter et al., 2017). For predicting farm size, we sampled from the posterior distribution and calculated the median (i.e., the geometric mean of the lognormal distribution) for the samples.

2.5. Random forest regression model

We also calibrated a random forest regression (RFR) to compare its farm size predictions to those of the Bayesian model. The RFR is a machine learning algorithm that operates on the principle of ensemble learning, by constructing many decision trees. Each tree is trained independently on a random subset of the data and provides its own prediction. The final prediction is then determined by aggregating (e.g., averaging) the predictions of all individual trees. We chose the RFR, as it is a widely used non-parametric machine learning technique that is easy to implement and excels in handling complex relationships between

many input features and the target variable. RFR models are robust in high dimensional feature-spaces and have very good predictive abilities across a variety of tasks (Breiman, 2001).

For our RFR, we used the same training sample, same predictors, and same pre-processing to compare the RFR against the Bayesian model. Additionally, we tested if the performance of the RFR improves by including all predictors that were excluded in the Bayesian model, for instance because of high correlation (Appendix 4). The hyperparameters for the RFR were chosen with a grid search on the number of trees, the maximum depth, and the maximum variables. We used 1,000 trees, a maximum of eight splits per tree, and six variables tried at each node. We performed all random forest regression modelling through the scikit-learn package in python (Pedregosa et al., 2011).

2.6. Accuracy assessment

To test the accuracy of both models, we calculated mean error (ME) (Eq. 14), root mean squared error (RMSE) (Eq. 15), mean absolute percentage error (MAPE) (Eq. 16) and precision (P) (Eq. 17). The ME provides insights whether our models systematically under- or over-estimate farm size. The RMSE represents how strongly our model estimates deviate from the observed data. MAPE is a relative error, indicating the proportion of the error relative to the true value. P is the statistical deviation around the true values corrected for the bias and is a measure of repeatability. A high precision means that the estimates are internally consistent, although they might come with consistent biases. The measures are calculated with the following formulas:

$$ME = \frac{1}{n} \sum_{i=0}^n \varepsilon_i \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n \varepsilon_i^2} \quad (15)$$

$$MAPE = \frac{1}{n} \sum_{i=0}^n \frac{\varepsilon_i}{r_i} \quad (16)$$

$$P = \sqrt{\frac{1}{n-1} \sum_{i=0}^n (\varepsilon_i - RMSE)^2} \quad (17)$$

where, n is the number of reference samples, r_i is the reference and ε_i is the difference between the estimation and r_i . We calculated all measures for the entire reference dataset, which included all observations that were not used for training. To test the model performance across farm sizes, we also stratified the reference dataset into non-equally sized bins of farm size (0–5 ha, 5–10 ha, 10–20 ha, 20–50 ha, 50–100 ha, 100–200 ha, 200–500 ha, 500–1,000 ha, 1,000–7,000 ha) and calculated the accuracy measures per farm size bin.

To explore the performance of the farm size estimation for different downstream applications, we further tested a spatial aggregation of the estimated farm sizes. For this, we calculated the mean of the estimated farm sizes for all fields within 15 km diameter hexagons. Note that this averages farm size estimates across all fields and does not reflect the average size of farms per hexagon, as this would require knowledge of the ownership of fields. We included this analysis, assuming that field-level estimation is subject to high uncertainties emerging from the multitude of factors influencing farm size, as described above, and therefore expecting better performance at aggregate spatial scales. We calculated the same error metrics in this experiment to be able to compare performance between field-level and spatially aggregated farm size estimation.

3. Results

3.1. Explaining farm size at the population-level

We found that field size and surrounding field sizes are both

positively associated with farm size (Fig. 4). Field size has a positive effect of 0.43 with a confidence interval between 0.35 and 0.51 (Table 4). This means that according to our model a 1 ha field is more likely part of a smaller farm than a 10-ha field. The effect of surrounding field size was also positive, but weaker (0.13) and with narrower confidence intervals (0.13–0.14). This means, that a 10-ha field surrounded by substantially larger fields is more likely part of a larger farm than a 10-ha field surrounded by substantially smaller fields. The effects of agricultural area per region and soil quality rating on farm size are negative but close to zero. Likewise, the interaction between surrounding field size and agricultural area per region has a slightly negative but close to zero effect. The interaction suggests that a given field with an area of, for example, 10 ha, which is surrounded by average-size fields and a moderate amount of agricultural area, is more likely to belong to a larger farm than a 10-ha field which is surrounded by large fields and a high amount of agricultural area.

3.2. Heterogeneous effects by crop type groups and Federal States

We found a clear distinction between East and West German states regarding the group specific variation from the population level intercept (Fig. 5A). East German states had a similar or higher intercept than the population level intercept, except for potatoes in Saxony-Anhalt and Brandenburg. West German states had a lower intercept than the population level with no exceptions.

We found that most East German states have positive deviations from the population level effect of field size (Fig. 5B), which translate into a stronger effect of field size on farm size. An exception is the “Other crops” group, for which we found slopes consistently below the population level and close to zero when the population level effect was included. This implies close to horizontal lines at the intercept, and hence, no association between field and farm size for this group. For sugar beet (except in Bavaria), cereals (except in Lower Saxony), and potato, the effect of field size on farm size was stronger than the crop-specific average. The effect of grassland was close to the population level effect or below, except for Thuringia. In general, the effect of sugar beet, potato, and cereals was stronger than for grassland and legumes. The variability in slopes was relatively high, with slopes ranging from negative slopes, e.g., for the “Other crops” group in Thuringia (mean effect of about -0.1) to slopes of about twice the value of the population level, e.g., for potato in Brandenburg (0.95) and Thuringia (0.92). A wider distribution of the model term, such as the sugar beet group in Thuringia, reflects a higher uncertainty in that estimate (e.g., due to small sample size in that group). Our model also included varying intercepts for the sigma parameter of our log normal likelihood function, for which the results can be found in Appendix 8.

The interplay between all parameters (e.g., the varying intercepts and slopes) is difficult to discern by merely assessing individual effects. Therefore, we used response curves stratified by the combinations of crop types and federal states for field size while keeping all other predictors at their respective means (Fig. 6). For almost all combinations, larger field size is associated with larger farm sizes except for the “Other crops” group, and for grassland in Bavaria. Large potato fields frequently belonged to large farms. For West German states, the model predicts much smaller farm sizes than for East German states. For Bavaria, our model predicts the smallest farms for most crop types. In Lower Saxony field size has a stronger effect on farm size and wider prediction intervals for all crop types. For the East German states, Thuringia has the strongest effect of field size on farm size and widest prediction intervals. This makes sense because we predict larger farms in Thuringia and the width of the log normal distribution increases with increasing mean even though sigma decreases with increasing field sizes (on the log scale we would see decreasing width of the intervals). In general, the prediction intervals are wide, indicating high uncertainties.

3.3. Error diagnoses for farm size predictions

We report size-specific performance statistics for the farm size predictions obtained from the Bayesian model and the two RFR models (Fig. 7, Appendix 9). The comparison of the mean error (ME), root mean squared error (RMSE), mean absolute percentage error (MAPE) and precision (P) across bins of farm size for the Bayesian model and the two RFR models (with selected predictors as in the Bayesian model, and with all predictors) revealed detailed insights on model performance and error distribution. We structure this section following a context-specific definition of small (0–50 ha), medium (50–500 ha), and large farms (500–7,000 ha).

Generally, and across all models, the size of small farms (0–50 ha) was overestimated, particularly for the smallest farm sizes. The ME for the Bayesian model reached 63 ha for the smallest farm size category (0–5 ha). However, the ME decreased to 38 ha for the largest category (20–50 ha) in this range. Logically, these moderate absolute errors translate into excessively high relative errors for very small classes. As such, the MAPE indicated significant discrepancies for farms of 0–5 ha, where the model predicted farm size with a MAPE of about 4,000% (0–5 ha, Bayesian model) or about 10,000% (0–5 ha, both RFR models). Note that this category also includes farms consisting of one field, for which we expect the model to predict a remaining farm size close to zero. Within the small farm sizes (0–50 ha), MAPE scores strongly decreased with increasing farm size, reaching the lowest values for farms of 20–50 ha with 125% (Bayesian model) to about 350% (RFR models).

The model predicts medium (50–500 ha) and large farm sizes (500–7,000 ha) reasonably well. Farms in these size classes represent about 66% of the fields contained in the data (Appendix 9). In this category, MAPE scores ranged from 50 to 70% (Bayesian model) and 40–114% (both RFR models). For the RFR models, MAPE scores remained below 81% for farm sizes above 100 ha. This tendency also translates into ME scores as low as 6 ha (Bayesian model) for farm sizes of 50–100 ha, which represents the largest category in the test data (287,596 fields). Similarly, a MAPE score of 50% was observed for farm sizes of 50–100 ha for the Bayesian model. The lowest MAPE scores of the RFR models were observed for farms of 500–1,000 ha (both 40%).

The ME were pronounced and negative for large farms (500–7000 ha), which revealed an underestimation in this category. For instance, large negative ME scores of $-1,156$ ha (RFR with all variables) and up to $-1,401$ ha (Bayesian model) were observed for the very large farms (1,000–7,000 ha). Relative errors in the large farms, however, were moderate with MAPE scores of 52% (500–1,000 ha) up to 68% (1,000–7,000 ha) for the Bayesian, and 40 (500–1,000 ha) to 55% (1,000–7,000 ha) for the RFR models.

Overall, the model uncertainties and precision (RMSE, P) were lower for the small and medium farm sizes (Bayesian model), or only in the large farm size categories (RFR models). RMSE and P scaled approximately with ME, indicating a relatively good performance in farm size categories below 200 ha but increasing rapidly for larger farms.

Regarding the overall model statistics, the Bayesian model's negative ME of -104 ha indicated an overall underestimation of farm sizes. The two RFR models had lower absolute ME (~ -7 ha) and thus only slightly underestimated farm sizes. The large overall RMSE of the Bayesian model (451 ha) and the RFR models (427 ha when using selected variables, and 423 ha, when using all variables) were comparable. The mean absolute percentage error (MAPE) revealed that farm size estimations were, on average, 250% off for the Bayesian model and 590% (selected variables), or 586% (all variables) off for the RFR models. This trend was mainly driven by the high MAPE scores for smaller farms, as described above. The high P scores of 439 ha for the Bayesian model and somewhat better scores for the RFR models with 427 ha (selected variables) and 423 ha (all variables) indicated inconsistent predictions. Note that all aggregate scores are highly sensitive to the outliers in the tails of the distribution, where small absolute deviations in the smallest farm size categories strongly penalize the MAPE scores, and the large absolute

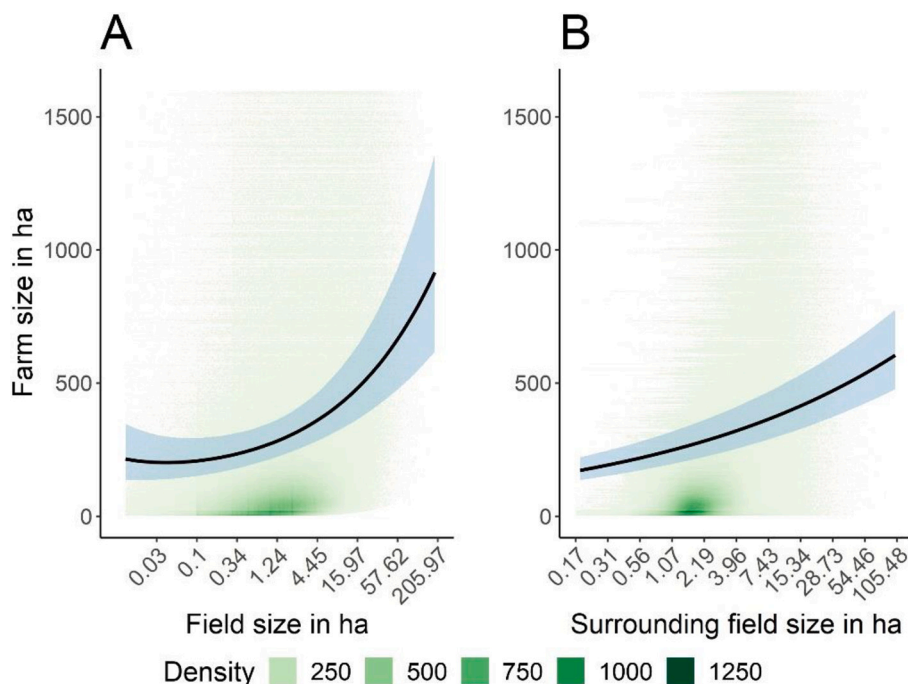


Fig. 4. Population level effect of (log) field size on farm size (left), and population level effect of surrounding field size on farm size (right). The axes are log-scaled, with labels representing values in hectares. The curves show the mean effect (black line) and 95% credible interval (blue shade) of the posterior distribution when all other model predictors are set to their mean value. Green shades depict density of observed data points. For the density the plane is divided into rectangles and the number of cases in each rectangle is counted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4

Model summary statistics. Parameters are summarized using mean (estimate) and the central 95% credible intervals. We report model summary statistics of the population-level and group-level effects. Population level effects represent the average relationship between predictors across the entire population (all crop type groups and all federal states), whereas group-level effects capture the variation that is specific to each group or cluster in the data. The group level effects describe the variation between groups, with higher values indicating more variation between groups. The varying slopes refer to the deviation from the population level effect of field size on farm size given federal state and crop type group.

Population-level Effects	Estimates	CI (95%)
Intercept (Eq. 3)	4.83	4.60–5.04
Sigma intercept (Eq. 10)	0.24	0.17–0.30
Field size (Eq. 4)	0.43	0.35–0.51
Surrounding field size (Eq. 4)	0.13	0.13–0.14
Agricultural area per region (Eq. 4)	−0.04	−0.04 to −0.03
Soil quality rating (Eq. 4)	−0.03	−0.03 – −0.03
Interaction: Surrounding field size - agricultural area per region (Eq. 4)	−0.02	−0.03 to −0.02
Sigma field size (Eq. 13)	−0.13	−0.13 to −0.13

Group-level Effects	Estimates	CI (95%)
Varying slopes (Eq. 8)	0.25	0.25–0.32
Varying intercepts (Eq. 6)	0.69	0.55–0.85
Varying intercepts sigma (Eq. 12)	0.21	0.17–0.26

deviations in the largest farm size categories have a strong effect on ME, RMSE, and P. Aggregate statistics thus mask the very heterogeneous performances of the models across the entire distribution of farm size and are not necessarily suitable to judge the performance of the approach as such.

To summarize: (i) in all models, medium (50–500 ha) and large (500–7,000 ha) farm sizes – which represent 66% of the data – were reasonably well predicted, depending on the model and the evaluation metric chosen (Fig. 7, Appendix 9); (ii) predictions in the small farm size category (0–50 ha) were inaccurate, and precision was low in the large farm sizes (1,000–7,000 ha), which reflected in high MAPE, and high RMSE and P values, respectively; and (iii) the RFR model incorporating all variables yielded the best overall performance statistics, but

differences in model performance were partly marginal and other models may outperform in specific farm size categories.

We further compared average farm size predictions with average reference farm sizes at an aggregated spatial level for the Bayesian model and the RFR with all variables (Fig. 8). For both models, MAPE, RMSE, and P improved substantially with spatial aggregation, with MAPE scores of 37% for both the Bayesian model and the RFR model. With regards to ME, the performance decreased from −104 ha at the field-level to −171.1 ha at the aggregate level for the Bayesian model, and from −7 ha at the field level to −35 ha at the aggregate level. The comparison revealed that both models captured the strong differences between the different states well, i.e., they predicted larger farm sizes for East Germany than for West Germany. However, both models strongly

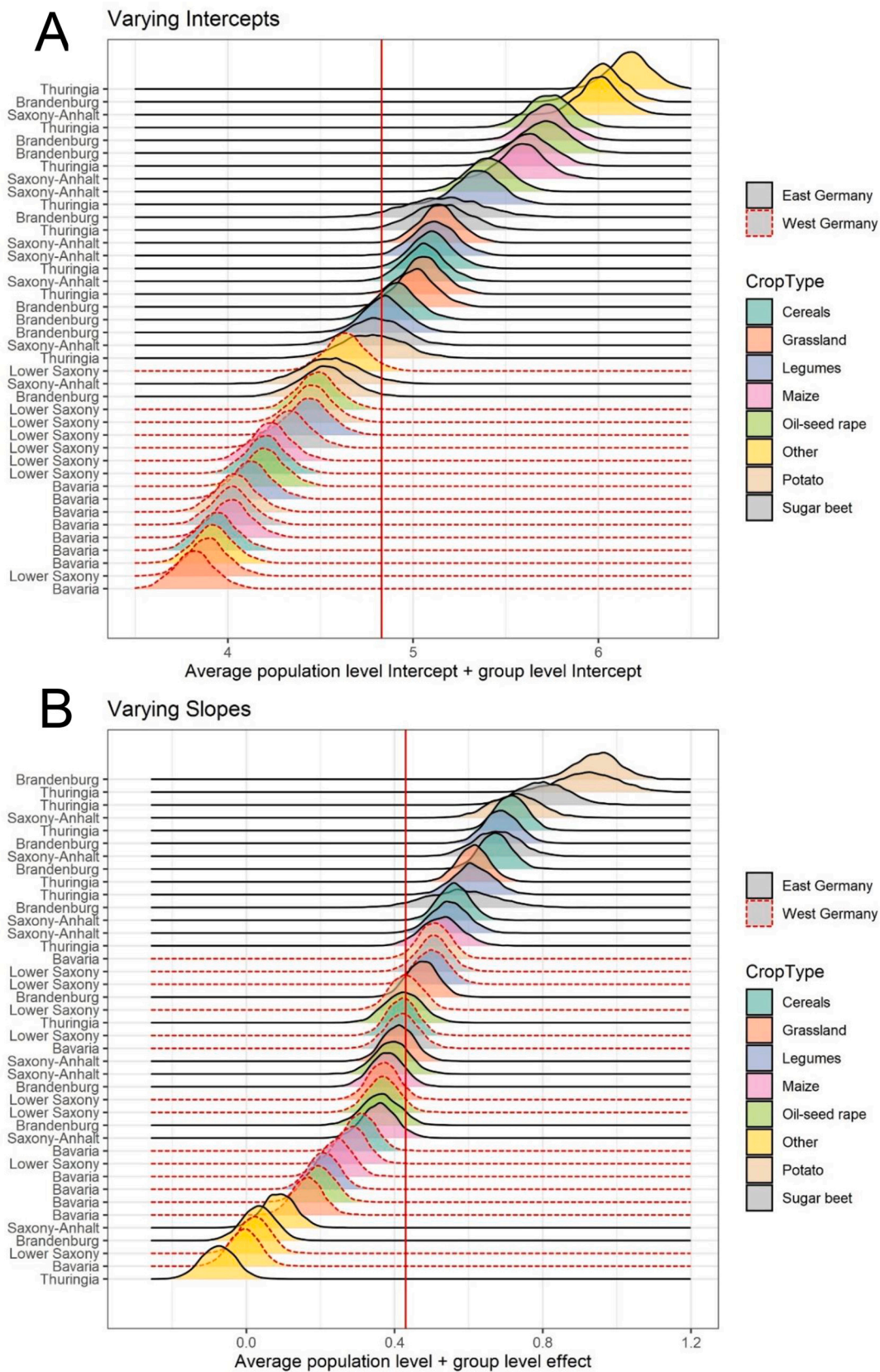


Fig. 5. A: Varying intercepts of the multilevel model for the location parameter μ . Values on the x-axis represent the group level effect added to the population level intercept. B: Varying slopes of log field size for the multilevel model. The x-axis shows the group-level slopes added to the population level effect. The colors represent crop type groups, black lines represent East German states and red dotted lines represent West German states. Vertical red lines indicate population level effect. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

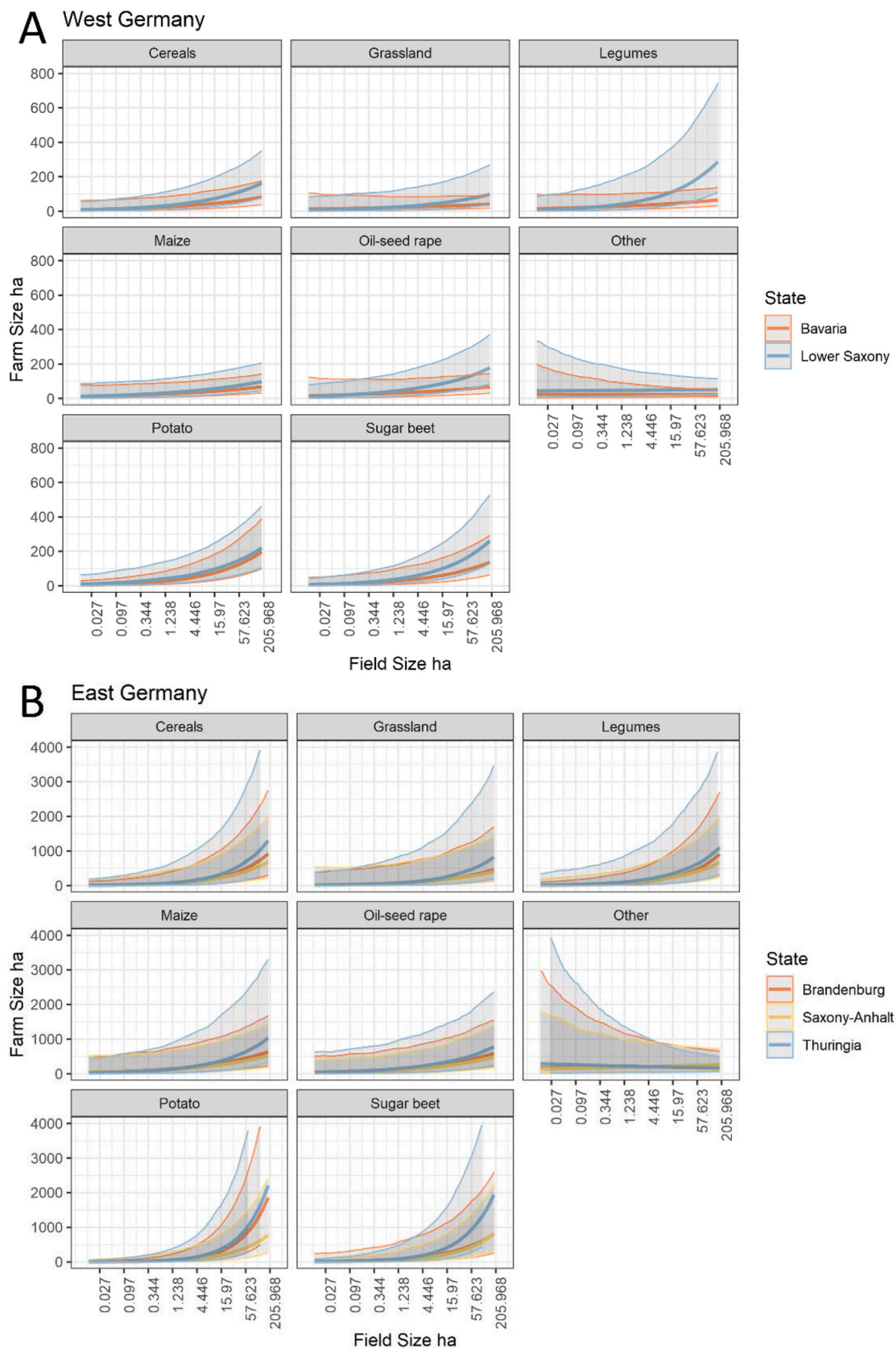


Fig. 6. A: Response curves for field size stratified by crop types for Bavaria and Lower Saxony. All other predictor variables are kept at their mean. We calculated the median posterior estimate and the posterior prediction intervals comprising both parameter uncertainty and predictive uncertainty. B: Response curves for field size stratified by crop types for Brandenburg and Saxony-Anhalt, and Thuringia. All other predictor variables are kept at their mean. We calculated the median posterior estimate and the posterior prediction intervals comprising both parameter uncertainty and predictive uncertainty. The axes are log-scaled, with labels representing values in hectares.

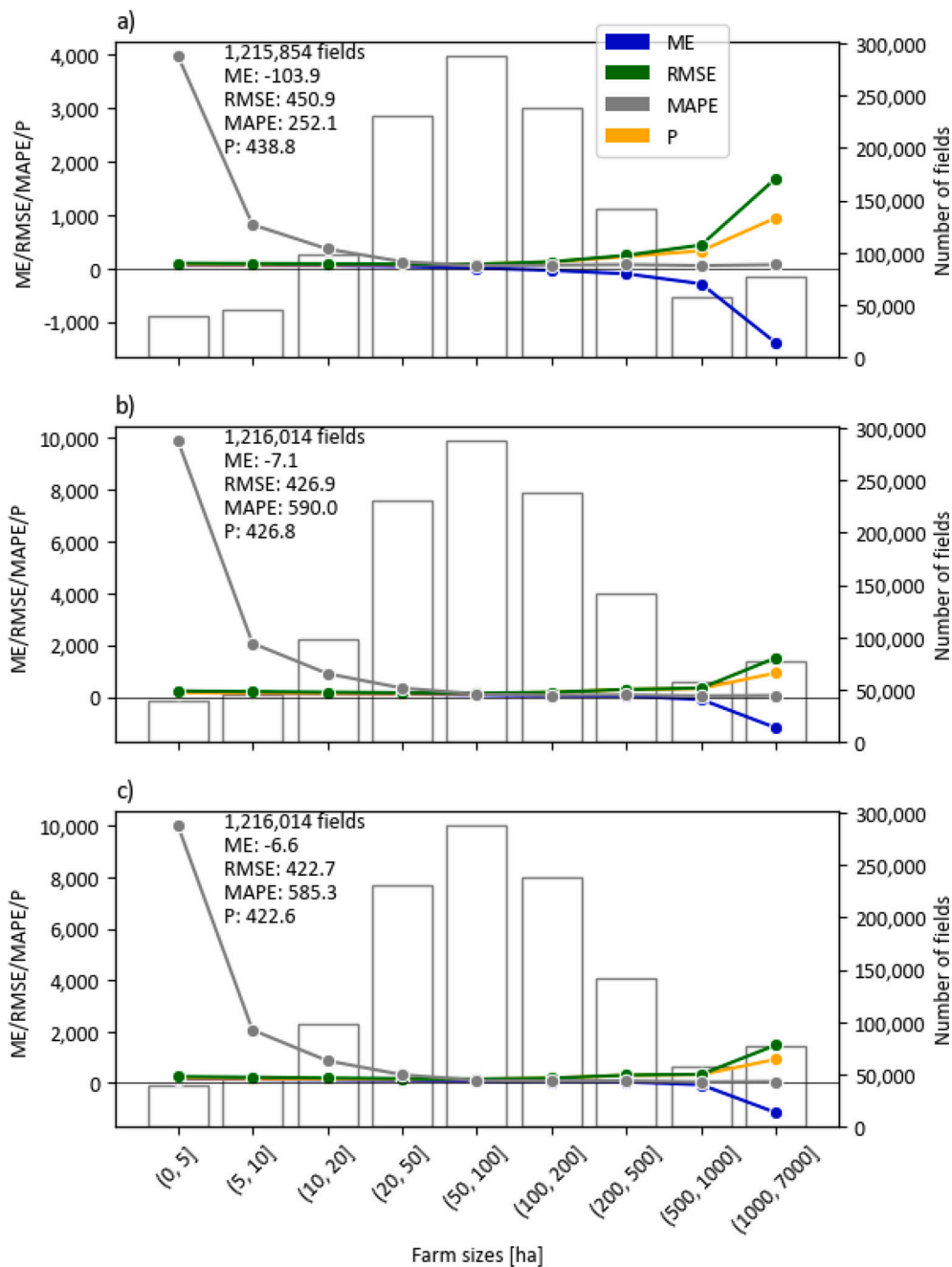


Fig. 7. ME, RMSE, MAPE and P for predictions by a) the Bayesian model, b) the random forest regression with the same observations and variables for training as the Bayesian model and, c) the random forest regression with the same observations but all available variables for training. Validation measurements are given across farm size categories (line plot) and across all observations (annotation). Bar plots indicate the number of fields per farm size category. A table with the exact values of all accuracy measures can be found in Appendix 9 and a table with summary statistics stratified by the farm size classes can be found in Appendix 10.

underestimated mean farm sizes above a certain threshold, which was approximately 700 ha for the Bayesian model and 1,000 ha for the RFR. As a result, spatial outliers of very large farms were not captured precisely in our predictions, particularly in the East German federal states. In the West German federal states, the RFR captured the intra-state variability better, e.g., it predicted the increasing farm sizes along the West-East gradient in Lower Saxony. Nevertheless, the exact locations of the spatial outliers were also incorrectly predicted by the RFR, which emphasizes that our models capture the general tendencies of farm size patterns, but not the local spatial variations.

4. Discussion

We analyzed the field size to farm size relationship (FFR) and its variation under different contextual settings in Germany. We modeled farm size based on field size and included variables on field (e.g., soil quality), administration (e.g., state), and site characteristics (e.g., proportion of agricultural land in a 1,000 m buffer). Our research objectives were motivated by the fact that spatially explicit knowledge about farm sizes is lacking for most regions, while information about field size can be derived from satellite imagery globally. Therefore, we constructed a Bayesian multilevel model and Random Forest Regression models to evaluate how effectively unknown farm sizes can be estimated based on readily available and spatially explicit data.

4.1. Understanding the FFR

Our results support the commonly assumed positive relationship between field size and farm size. Farms generally aim to create fields that match farm-level access to labor and machinery, which often means that field sizes increase along with capitalization and mechanization of farming. Larger consolidated fields can increase profitability and efficiency by benefiting from economies of scale (Clough et al., 2020; Heinrichs et al., 2021; Latruffe and Piet, 2014) due to optimized use of machinery. Given the limited financial resources of smaller farms, the possibilities of creating and managing large fields may thus be limited. We suspect that a mutual causation might occur. Farms with larger fields may have an economic advantage and thus be able to increase in size, and conversely, farms of growing sizes strive for consolidating larger fields to be able to benefit from economies of scale. Further, larger farms would focus more on specific field crops (Weigel et al., 2018), for which economies of scale can be achieved, which likely leads to larger field sizes. Our results suggest that this is especially the case for sugar beet and potato as large fields with these crops more likely belong to larger farms than other fields with a similar size but different crop.

However, our results underline the multifactorial and often complex nature of the FFR. In East Germany, for instance, the relationship between farm and field size is stronger than in West Germany, and it is weakest in Bavaria, where farms and fields are smallest. The observed difference between East and West Germany is a political legacy, as in former times, the government of the former German Democratic

Republic implemented land consolidation programs, while western Germany promoted smaller family-owned farms (Noack et al., 2022). Our findings highlight that the implementation of territorial policies likely has a strong effect on the FFR, to a degree that outweighs within-state effects related to the biophysical context of agricultural land, e.g. soil quality, or landscape configurations, e.g. agricultural area per region or surrounding field sizes.

Furthermore, we found that the FFR varied considerably between crop type groups. For instance, our model suggested that a 100-ha potato field in Thuringia is more likely to belong to a larger farm than a 100-ha maize field. This may suggest that large potato fields are economically rather viable for large farms, while large maize fields can be managed by smaller farms, too. For many crop groups (cereals, grassland, legumes, maize, potatoes, and sugar beet), the FFR in East Germany was stronger than in West Germany. Only for rapeseed, the relationship was stronger in West German, as compared to East German states. In addition, for the crop type group “other”, we observed no relationship between field size and farm size, and wide predictions intervals for smaller field sizes. This means that for this group, field sizes do not necessarily increase with farm sizes, and that both small or large farms may operate crops contained in the group “other” on small fields. This is likely because the crop type group consists of very diverse crops. On the one hand, it includes labor-intensive crops, such as vegetables, which are more difficult to mechanize and mostly grown by smaller farms (Joshi et al., 2006; Weigel et al., 2018; Wolz, 2013). On the other hand, this crop type group includes permanent crops, for which we find very large fields in our

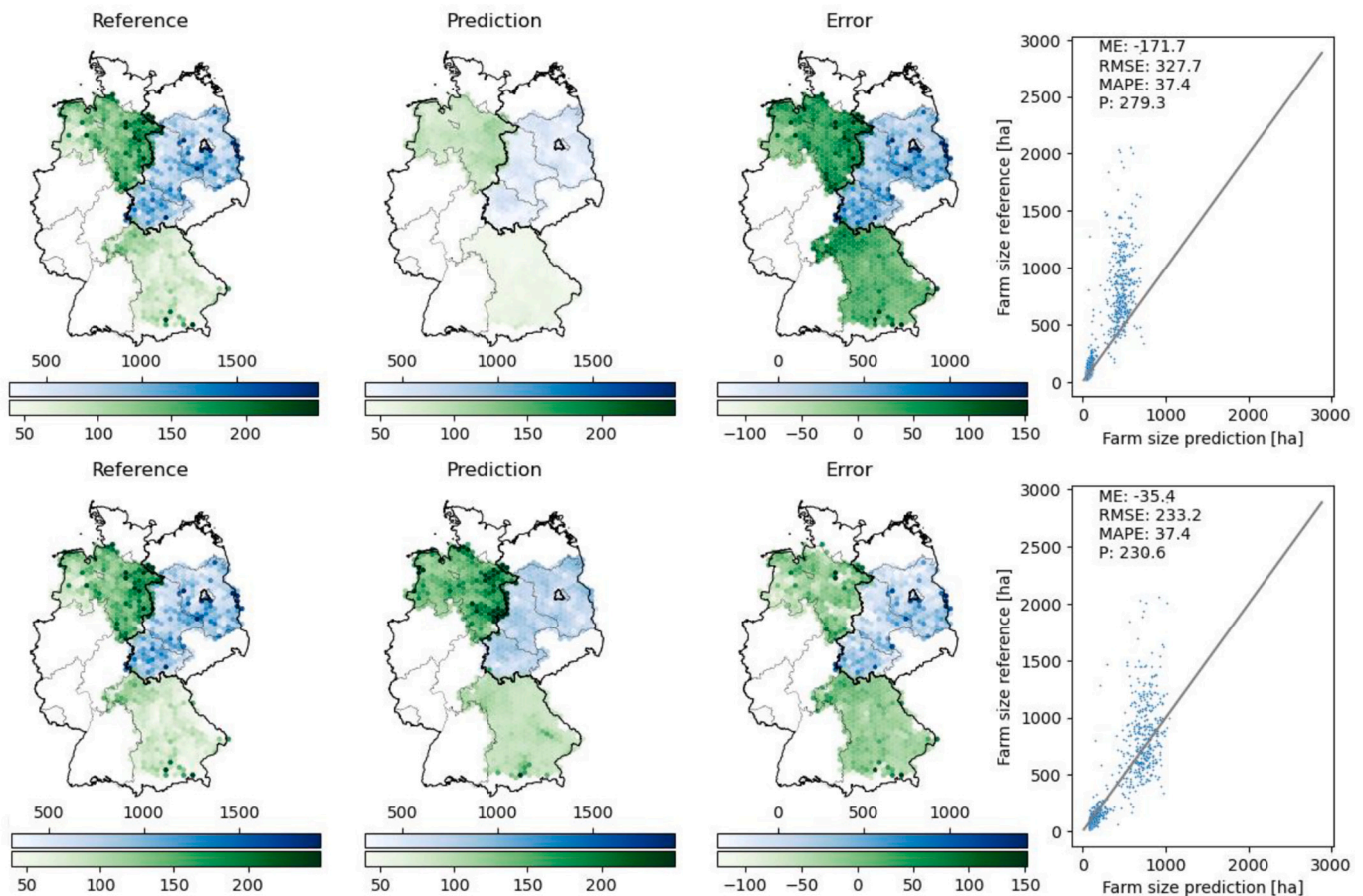


Fig. 8. Model results spatially aggregated to hexagons with a vertical and a horizontal diameter of 15 km ($\sim 195 \text{ km}^2$) for the a) Bayesian model (top row) and b) the random forest regression with the same observations but all available variables for training (bottom row). The aggregated values in the hexagons represent the mean reference farm size, mean predicted farm size, and mean error values across all fields that fall into the hexagon. The columns show observed farm size [ha], predicted farm size [ha], prediction errors [ha] and scatterplots of reference farm sizes vs. predicted farm sizes with ME, RMSE, MAPE and P. The green scale represents the range of farm sizes in West Germany, while the blue scale represents the range of farm size in East Germany. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

study area. Overall, the FFR is thus dependent not only on crop types, but also the context in which a particular crop is grown.

As hypothesized, the size of surrounding fields was important in predicting farm size. Large fields surrounded by other large fields likely belong to a large farm. In contrast, a large field surrounded by small fields is likely part of a smaller farm. The very small effect of soil quality on farm size was unexpected. However, as we only included soil quality on the population level, one explanation is that multiple heterogeneous effects cancel each other out. These could be that best soils tend to be acquired by larger farms with high capital endowments. On the other hand, larger areas of low-quality soil are required to achieve similar production levels, which might lead to an association of low-quality soil and large farms. The agricultural area per region has a small negative effect on farm size. Multiple co-existing and context-dependent hypotheses support heterogeneous effects that could - in sum - cancel each other. Regions with abundant agricultural lands would typically provide various forms of agglomeration economies to farmers (such as higher density of services, skilled labor, etc.), potentially allowing farms to be more competitive and consolidate over time. Conversely, prime agricultural lands, sustaining higher yields, may have been leading over time to higher population densities, which in turn may result in landscapes occupied by many small farms (Ellis, 2021).

4.2. Estimating farm size

We tested a Bayesian model and RFR to predict farm size. Our models showed broad predictive power for the bulk of farm size categories but revealed a high variability in the FFR and the associated predictive uncertainties for the smallest and largest farm size categories. On the one hand, the size of very small farms in our data is often overpredicted. This could arise from the fact that although large farms tend to have larger fields, they also cultivate fields that are just as small as those found on small farms, leading to higher uncertainty for predicting farm size for small fields. On the other hand, the sizes of large farms are underpredicted. Our results imply that farm size increases with field sizes, but when field size approaches the upper limit of a manageable land unit, the relationship between field and farm size fades. Empirically, our aggregated predictions at the 15 km hexagon level revealed that this is the case when average farm size is larger than 700–1,000 ha (compare scatterplots of Fig. 8). Although we observed high uncertainties at both ends of the farm size gradient, we find that for medium-sized (50–500 ha) and large farms (500–7,000 ha), which make up the bulk of all farms (66%), the uncertainties are reasonably low with MAPE scores of 50–70% (Bayesian model) or 40–114% (RFR models). The lower performance of the models for the smallest and (MAPE) somewhat also the largest farm sizes (RMSE, P) could partly be due to a lower sample size for these farm classes compared to the medium range. Analyzing the FFR across heterogeneous geographical contexts, including, for instance, smallholder dominated regions, might shed light on this aspect.

For the Bayesian model we assumed linear relationships between explanatory variables, while the RFR can capture non-linear relationships. The comparison of the models revealed that overall, the RFR performs slightly better in predicting farm size (based on ME and RMSE, but at the expense of severe overpredictions on the smallest farm size categories), especially when we included variables that had to be excluded from the Bayesian model due to multicollinearity. The better performance of the RFR might indicate that some of the relationships are nonlinear. Importantly, although the RFR performed better, predictions also came with high errors and moderate accuracy (Fig. 7).

The spatial patterns of the predictions displayed smooth gradients, which failed to capture hot spots of very high or low farm sizes. Spatially aggregating farm size predictions improved the accuracy, with overall MAPE scores of 37% for both models, as the influence of local outliers decreases. The poor performance in capturing positive outliers of farm size likely relates to the fact that field sizes have a natural size limit at a

certain threshold (as discussed above), which may also explain why in some cases, the prediction of large farm sizes was spatially close to observed farm sizes, but not precisely matching, as for instance observed in the RFR results in southern Bavaria (Fig. 7).

4.3. Missing variables and limitations

We used federal states to account for differences in policies and historical backgrounds but likely missed out on important information, which may play out below the level of federal states. For example, the presence of land consolidation programs used to consolidate ownership patterns may impact field sizes positively (Sklenicka et al., 2009). The existence and the success of consolidation programs might vary strongly between and within the federal states. Further, succession rules and historic developments, such as land expropriations and restitutions, can differ strongly between regions, which still influences ownership patterns today (Gatterer et al., 2024; Jänicke and Müller, 2024). Therefore, including spatially more detailed information on such processes might explain more of the local variations that we were not able to capture.

Although we incorporated information on the crop type groups cultivated per field, this variable only captures a snapshot in time, as farmers plant different crops over several years. Therefore, the response of field size per crop type might also change over time. Using a dataset that spans multiple years and potentially includes information on crop sequences (Jänicke et al., 2022) could help to mitigate this uncertainty. However, farmers are unlikely to cultivate a given crop in a field that is (from a management perspective) too small or too large.

Additionally, we could not include information on farm specialization or total economic farm size, as we used field sizes to predict farm sizes and not vice versa. For example, there is evidence showing that organic farms have significantly smaller fields than conventional farms (Norton et al., 2009). Further, Levin (2006) hypothesized that there might be a difference between part-time and full-time farmers, as the latter are more interested in production maximization, and thus more likely to increase the size of their fields. These variables are likely to carry additional explanatory power for the FFR, however, we could not include this information given the setup of our analysis. Studying the reverse relationship, i.e., explaining field sizes with farm sizes could allow including this information and could thereby provide additional insights.

Adding other contextual variables such as on climate could help to improve the model further for larger and climatically more heterogeneous study areas. Further, Levin (2006) hypothesized that the influence of recreational, scenic, and environmental interests in areas with a large urban population explains lower field sizes and higher densities of field divides and small biotopes. Deininger and Byerlee (2012) also highlighted several key factors driving large fields and farms related to their precise location, such as availability of infrastructure and technology, which could be tested in the future.

Lastly, our investigation of the field size to farm size relationship builds on the assumption that field size data is abundant. We had the opportunity to explore the FFR in unprecedented conditions due to the availability of IACS data. We argue that in the more common setting where similar data is not available, field sizes could be estimated from satellite imagery. Recent developments in EO and computer vision present emerging opportunities to produce spatially detailed and timely field size estimates across large areas (Waldner and Diakogiannis, 2020). EO-based field delineations are Essential Agricultural Variables (EAV) that facilitate monitoring the progress towards the Sustainable Development Goals (SDG) (Whitcraft et al., 2019), as they are prerequisites for field-level assessments of productivity and land management (Lobell et al., 2020). EO-based field delineation has nearly reached operational status in consolidated landscapes with large fields (Waldner et al., 2021). Data bottlenecks continue to challenge field delineation in smallholder systems in low- and middle-income countries (Nakalembe and Kerner, 2023; Persello et al., 2019; Wang et al., 2022), but

pioneering national scale datasets of individual fields are emerging (Estes et al., 2022). The continued development of satellite-based field size estimates and land use indicators opens new opportunities to enhance our knowledge about farm size distributions for large regions. However, EO-based field delineation is often geometrically imperfect (Tetteh et al., 2023), and the resulting field size estimates are subject to additional uncertainty that arise, for example, when the same crop type is planted on two neighboring fields or when trees are part of the farming system. We highlight that in-depth error diagnosis of EO-based field size estimates is needed to evaluate the suitability of field delineations for the respective downstream applications (Rufin et al., 2023; Wang et al., 2022). While this was beyond the scope of our study, we encourage further research on how to account for such uncertainties.

4.4. Policy implications

The findings of the study on the relationship between field size and farm size (FFR) in Germany, leveraging Bayesian multilevel and machine learning regression models, carry significant implications for both policy and practice in the agricultural sector. Firstly, the ability to estimate farm sizes from field size and other readily available predictors marks a substantial advancement in agricultural data analysis, although uncertainties for very small and large farms remain high. Our approach can be used for a more nuanced monitoring of farm dynamics, enabling policymakers to design interventions that are better tailored to the needs of diverse agricultural systems in circumstances where farm size data are not available.

For policy development, our analysis highlights the potential of integrating Earth observation data into agricultural planning and monitoring frameworks. Better knowing the FFR and its heterogeneity can help to target policies, acknowledging that environmental impacts are more directly linked to field sizes (such as landscape heterogeneity, edges, and mosaics; Sirami et al., 2019; Tschardt et al., 2021; Wesemeyer et al., 2023), while socio-economic impacts are more strongly related to farm size (Noack and Larsen, 2019). For example, interventions aiming at sustainable land use and optimizing farm productivity could be adjusted according to the specific patterns of farm sizes within regions. This approach would ensure that policies are not only more targeted but also more responsive to the spatial and temporal dynamics of farm sizes.

Our study was situated in a uniquely data-rich environment which facilitated the exploration of statistical relationships based on large amounts of data. The relevance of our findings underpins the added value that can result from a public release of the European IACS data, or similar datasets around the globe. Yet unfortunately, such datasets are scarce in other world regions. Consequently, further establishing robust statistical relationships between field size, contextual predictor variables, and farm size would allow for broader monitoring capabilities in these regions. Advancing our understanding of the FFR opens new pathways for complementing traditional farm surveys and censuses, which are often limited by infrequent updates and reporting for aggregate spatial units. By incorporating spatially explicit data derived from Earth observation, agricultural statistics could be enriched, offering a more detailed and up-to-date picture of agricultural landscapes. This could enhance the design and implementation of agricultural policies, ensuring they are grounded in the current realities of farm management and land use.

5. Conclusion

This study presents the first empirical analysis of the field size to farm size relationship (FFR) and its variation in a data-rich setting in Germany. Our work yielded four key insights. First, we found that farm size increases with field size, but predictions of farm size are associated with high uncertainty and heteroscedasticity. The relationship between farm and field size is stronger and the predictions were more accurate for

medium and large farms in the range of 50–7,000 ha, which represent the bulk of the farms in the study region. Second, predictions tended to be more accurate when aggregated to larger regions, e.g., for landscape units. Third, no clear relationship between field and farm sizes exists for some crop type groups, such as vegetables or flowers. Fourth, machine learning models have better overall performance than the Bayesian model for predicting farm sizes, while the Bayesian model is more explanatory and allows estimating prediction uncertainty. In summary, we show that field sizes and farm sizes are indeed related, but more research is needed to assess the transferability of our results beyond the context of Germany. Also, a deeper understanding or theory about how and why larger fields are associated with larger farms is needed.

Our study provides a valuable foundation for developing more informed and effective agricultural policies. By leveraging spatially explicit data and advanced modelling techniques, policymakers can foster more sustainable, productive, and equitable agricultural systems. The approach outlined in the study offers a promising direction for future research and policy development, emphasizing the need for continuous innovation in agricultural data analysis and policy design.

CRedit authorship contribution statement

Clemens Jänicke: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Maximilian Wesemeyer:** Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Cristina Chiarella:** Writing – review & editing. **Tobia Lakes:** Writing – review & editing, Resources, Funding acquisition. **Christian Levers:** Writing – review & editing. **Patrick Meyfroidt:** Writing – review & editing, Funding acquisition. **Daniel Müller:** Writing – review & editing, Resources, Funding acquisition. **Marie Pratzner:** Writing – review & editing, Methodology. **Philippe Rufin:** Writing – original draft, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

PR and PM acknowledge support by the F.R.S.-FNRS under grant no. T.0154.21, grant no. 1.B422.24. and grant no. 1.B.264.22F. CL gratefully acknowledges support by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement 796451 (FFSize). CJ and DM gratefully acknowledge support by the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101081307. We thank the Ministry of Agriculture, Environment, and Climate Protection (MLUK) of the Federal State of Brandenburg, the Ministry for Environment, Agriculture, and Energy (MULE) of Saxony-Anhalt, the State Ministry for Nutrition, Agriculture, and Forests of Bavaria, the Ministry for Human Nutrition, Agriculture and Consumer Protection of Lower Saxony, and the Thuringian State Office for Agriculture and Rural Areas (TLLLR) for providing the IACS data. This work was supported by the Deutsche Forschungsgemeinschaft as part of Research Unit 2569, 'Agricultural Land Markets—Efficiency and Regulation' and contributes to the Global Land Programme (<https://glp.earth>). We gratefully acknowledge the data provided by the Ministry of Economy, Tourism, Agriculture, and Forestry of the Federal State Saxony-Anhalt for the project "Market power in agricultural land markets - significance,

measurement, delimitation,” funded by the Federal Office for Agriculture and Food (BLE). Lastly, we are grateful to two anonymous reviewers for their thoughtful and constructive comments that helped to improve this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.agry.2024.104088>.

References

- (C) BGR, 2013. Soil quality rating for cropland in Germany 1:1,000,000 (1.0). <https://gd.k.gdi-de.org/geonetwork/srv/api/records/3DBDC1EE-81E9-41A2-916E-1281DD6C7A8>.
- Bevis, L.E.M., Barrett, C.B., 2020. Close to the edge: high productivity at plot peripheries and the inverse size-productivity relationship. *Journal of Development Economics* 143, 102377. <https://doi.org/10.1016/j.jdeveco.2019.102377>.
- Bey, A., Jetimane, J., Lisboa, S.N., Ribeiro, N., Sitoe, A., Meyfroidt, P., 2020. Mapping smallholder and large-scale cropland dynamics with a flexible classification system and pixel-based composites in an emerging frontier of Mozambique. *Remote Sens. Environ.* 239, 111611 <https://doi.org/10.1016/j.rse.2019.111611>.
- Branthomme, A., Merle, C., Kindgard, A., Lourenço, A., Ng, W., D'Annunzio, R., Shapiro, A., 2023. How Much Do Large-Scale and Small-Scale Farming Contribute to Global Deforestation. FAO, Rome, Italy.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Bürkner, P.-C., 2018. Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal* 10, 395–411. <https://doi.org/10.32614/RJ-2018-017>.
- Carletto, C., Savastano, S., Zezza, A., 2013. Fact or artifact: the impact of measurement errors on the farm size–productivity relationship. *J. Dev. Econ.* 103, 254–261. <https://doi.org/10.1016/j.jdeveco.2013.03.004>.
- Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A., 2017. Stan: A probabilistic programming language. *J. Stat. Softw.* 76 (1), 1–32. <https://doi.org/10.18637/jss.v076.i01>.
- Chiarella, C., Meyfroidt, P., Abeygunawardane, D., Conforti, P., 2023. Balancing the trade-offs between land productivity, labor productivity and labor intensity. *Ambio* 52 (10), 1618–1634. <https://doi.org/10.1007/s13280-023-01887-4>.
- Clough, Y., Kirchweiger, S., Kantelhardt, J., 2020. Field sizes and the future of farmland biodiversity in European landscapes. *Conserv. Lett.* 13 (6), e12752 <https://doi.org/10.1111/conl.12752>.
- Dang, D.K., Patterson, A.C., Carrasco, L.R., 2019. An analysis of the spatial association between deforestation and agricultural field sizes in the tropics and subtropics. *PLoS One* 14 (1), e0209918.
- Dannenberg, P., Kuemmerle, T., 2010. Farm size and land use pattern changes in Post-socialist Poland*. *Prof. Geogr.* 62 (2), 197–210. <https://doi.org/10.1080/00330120903546312>.
- De Groote, H., Traoré, O., 2005. The cost of accuracy in crop area estimation. *Agr. Syst.* 84 (1), 21–38. <https://doi.org/10.1016/j.agry.2004.06.008>.
- Debonne, N., Bürgi, M., Diogo, V., Helfenstein, J., Herzog, F., Levers, C., Mohr, F., Swart, R., Verburg, P., 2022. The geography of megatrends affecting European agriculture. *Glob. Environ. Chang.* 75, 102551 <https://doi.org/10.1016/j.gloenvcha.2022.102551>.
- Deininger, K., Byerlee, D., 2012. The rise of large farms in land abundant countries: do they have a future? *World Dev.* 40 (4), 701–714. <https://doi.org/10.1016/j.worlddev.2011.04.030>.
- Ellis, E.C., 2021. Land use and ecological change: A 12,000-year history. In: *Annual Review of Environment and Resources*, vol. 46, pp. 1–33. <https://doi.org/10.1146/annurev-environ-012220-010822>.
- Estes, L.D., Ye, S., Song, L., Luo, B., Eastman, J.R., Meng, Z., Zhang, Q., McRitchie, D., Debats, S.R., Muhando, J., 2022. High resolution, annual maps of field boundaries for smallholder-dominated croplands at national scales. *Frontiers in Artificial Intelligence* 4, 744863.
- European Environment Agency (EEA), 2016. European Digital Elevation Model [Dataset]. <https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1?tab=metadata>.
- Eurostat, P., 2022. Farms and Farmland in the European Union—Statistics.
- Gatterer, M., Leonhardt, H., Salhofer, K., Morawetz, U., 2024. The legacy of partible inheritance on farmland fragmentation: evidence from Austria. *Land Use Policy* 140, 107110. <https://doi.org/10.1016/j.landusepol.2024.107110>.
- 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. *Agr. Syst.* 187, 102992 <https://doi.org/10.1016/j.agry.2020.102992>.
- Helfand, S.M., Taylor, M.P.H., 2021. The inverse relationship between farm size and productivity: refocusing the debate. *Food Policy* 99, 101977. <https://doi.org/10.1016/j.foodpol.2020.101977>.
- Jänicke, C., Müller, D., 2024. Revealing agricultural land ownership concentration with cadastral and company network data. *Agric. Hum. Values.* <https://doi.org/10.1007/s10460-024-10590-3>.
- Jänicke, C., Goddard, A., Stein, S., Steinmann, H.-H., Lakes, T., Nendel, C., Müller, D., 2022. Field-level land-use data reveal heterogeneous crop sequences with distinct regional differences in Germany. *Eur. J. Agron.* 141, 126632 <https://doi.org/10.1016/j.eja.2022.126632>.
- Jayne, T.S., Wineman, A., Chamberlin, J., Muyanga, M., Yeboah, F.K., 2022. Changing farm size distributions and agricultural transformation in sub-Saharan Africa. *Ann. Rev. Resour. Econ.* 14 (1), 109–130. <https://doi.org/10.1146/annurev-resource-111220-025657>.
- Joshi, P., Joshi, L., BIRTHAL, P.S., 2006. Diversification and its impact on smallholders: evidence from a study on vegetable production. *Agric. Econ. Res. Rev.* 19 (347–2016–16776), 219–236.
- Kuemmerle, T., Erb, K., Meyfroidt, P., Müller, D., Verburg, P.H., Estel, S., Haberl, H., Hostert, P., Jepsen, M.R., Kastner, T., Levers, C., Lindner, M., Plutzer, C., Verkerk, P. J., Zanden, E.H., Reenberg, A., 2013. Challenges and opportunities in mapping land use intensity globally. *Curr. Opin. Environ. Sustain.* 5 (5), 484–493. <https://doi.org/10.1016/j.cosust.2013.06.002>.
- Latruffe, L., Piet, L., 2014. Does land fragmentation affect farm performance? A case study from Brittany, France. *Agr. Syst.* 129, 68–80. <https://doi.org/10.1016/j.agry.2014.05.005>.
- Leonhardt, H., Hüttl, S., Lakes, T., Wesemeyer, M., Wolff, S., 2023. Use cases of the integrated administration and control system's plot-level data: protocol and pilot analysis for a systematic mapping review. *German Journal of Agricultural Economics* 72 (3/4), 168–184. <https://doi.org/10.30430/gjae.2023.0385>.
- Levin, G., 2006. Farm size and landscape composition in relation to landscape changes in Denmark. *Geografisk Tidsskrift-Danish Journal of Geography* 106 (2), 45–59. <https://doi.org/10.1080/00167223.2006.10649556>.
- Liebert, J., Benner, R., Bezner Kerr, R., Björkman, T., De Master, K.T., Gennet, S., Gómez, M.L., Hart, A.K., Kremen, C., Power, A.G., Ryan, M.R., 2022. Farm size affects the use of agroecological practices on organic farms in the United States. *Nature Plants* 8 (8), 897–905. <https://doi.org/10.1038/s41477-022-01191-1>.
- Lobell, D.B., Azzari, G., Burke, M., Gourlay, S., Jin, Z., Kilic, T., Murray, S., 2020. Eyes in the sky, boots on the ground: assessing satellite- and ground-based approaches to crop yield measurement and analysis. *Am. J. Agric. Econ.* 102 (1), 202–219. <https://doi.org/10.1093/ajae/aaz051>.
- Lowder, S.K., Skoet, J., Raney, T., 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Dev.* 87, 16–29. <https://doi.org/10.1016/j.worlddev.2015.10.041>.
- Meissner, L., Musshoff, O., 2022. Revisiting the relationship between farmland prices and soil quality. *Q. Open* 2 (2), qoac017. <https://doi.org/10.1093/qopen/qoac017>.
- Meyfroidt, P., 2017. Mapping farm size globally: benchmarking the smallholders debate. *Environ. Res. Lett.* 12 (3), 031002 <https://doi.org/10.1088/1748-9326/aa5ef6>.
- Nakalembe, C., Kerner, H., 2023. Considerations for AI-EO for agriculture in sub-Saharan Africa. *Environ. Res. Lett.* 18 (4), 041002.
- Noack, F., Larsen, A., 2019. The contrasting effects of farm size on farm incomes and food production. *Environ. Res. Lett.* 14 (8), 084024 <https://doi.org/10.1088/1748-9326/ab2dbf>.
- Noack, F., Larsen, A., Kamp, J., Levers, C., 2022. A bird's eye view of farm size and biodiversity: the ecological legacy of the iron curtain. *Am. J. Agric. Econ.* 104 (4), 1460–1484. <https://doi.org/10.1111/ajae.12274>.
- Norton, L., Johnson, P., Joys, A., Stuart, R., Chamberlain, D., Feber, R., Firbank, L., Manley, W., Wolfe, M., Hart, B., Mathews, F., Macdonald, D., Fuller, R.J., 2009. Consequences of organic and non-organic farming practices for field, farm and landscape complexity. *Agr. Ecosyst Environ.* 129 (1), 221–227. <https://doi.org/10.1016/j.agee.2008.09.002>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12 (85), 2825–2830.
- Pe'er, G., Dicks, L.V., Visconti, P., Arlettaz, R., Baldi, A., Benton, T.G., Collins, S., Dieterich, M., Gregory, R.D., Hartig, F., Henle, K., Hobson, P.R., Kleijn, D., Neumann, R.K., Robijns, T., Schmidt, J., Schwartz, A., Sutherland, W.J., Turbé, A., et al., 2014. EU agricultural reform fails on biodiversity. *Science* 344 (6188), 1090–1092. <https://doi.org/10.1126/science.1253425>.
- Persello, C., Tolpekin, V.A., Bergado, J.R., de By, R.A., 2019. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote Sens. Environ.* 231, 111253 <https://doi.org/10.1016/j.rse.2019.111253>.
- Ricciardi, V., Mehrabi, Z., Wittman, H., James, D., Ramankutty, N., 2021. Higher yields and more biodiversity on smaller farms. *Nature Sustainability* 4 (7), 651–657. <https://doi.org/10.1038/s41893-021-00699-2>.
- Roschewitz, I., Thies, C., Tscharntke, T., 2005. Are landscape complexity and farm specialisation related to land-use intensity of annual crop fields? *Agr. Ecosyst Environ.* 105 (1–2), 87–99.
- Rufin, P., Wang, S., Lisboa, S.N., Hemmerling, J., Tulbure, M.G., Meyfroidt, P., 2023. Taking it further: leveraging pseudo labels for field delineation across label-scarce smallholder regions. *arXiv preprint arXiv:2312.08384*.
- Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M., West, P.C., 2016. Subnational distribution of average farm size and smallholder contributions to global food production. *Environ. Res. Lett.* 11 (12), 124010 <https://doi.org/10.1088/1748-9326/11/12/124010>.
- Scown, M.W., Brady, M.V., Nicholas, K.A., 2020. Billions in misspent EU agricultural subsidies could support the sustainable development goals. *One Earth* 3 (2), 237–250. <https://doi.org/10.1016/j.oneear.2020.07.011>.
- Sikor, T., Müller, D., Stahl, J., 2009. Land fragmentation and cropland abandonment in Albania: implications for the roles of state and Community in Post-Socialist Land Consolidation. *The Limits of State-Led Land Reform* 37 (8), 1411–1423. <https://doi.org/10.1016/j.worlddev.2008.08.013>.

- Sirami, C., Gross, N., Baillod, A.B., Bertrand, C., Carrié, R., Hass, A., Henckel, L., Miguet, P., Vuillot, C., Alignier, A., Girard, J., Batáry, P., Clough, Y., Violle, C., Giralt, D., Bota, G., Badenhausser, I., Lefebvre, G., Gauffre, B., Fahrig, L., 2019. Increasing crop heterogeneity enhances multitrophic diversity across agricultural regions. *Proc. Natl. Acad. Sci.* 116 (33), 16442–16447. <https://doi.org/10.1073/pnas.1906419116>.
- Sklenicka, P., Hladík, J., Štřelček, F., Kottova, B., Lososová, J., Číhal, L., Šálek, M., 2009. Historical, environmental and socio-economic driving forces on land ownership fragmentation, the land consolidation effect and the project costs. *Agricultural Economics (AGRICECON)* 55, 571–582. <https://doi.org/10.17221/115/2009-AGRICECON>.
- Tetteh, G.O., Schwieder, M., Erasmí, S., Conrad, C., Gocht, A., 2023. Comparison of an optimised multiresolution segmentation approach with deep neural networks for delineating agricultural fields from Sentinel-2 images. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 91 (4), 295–312. <https://doi.org/10.1007/s41064-023-00247-x>.
- Tscharntke, T., Grass, I., Wanger, T.C., Westphal, C., Batáry, P., 2021. Beyond organic farming – harnessing biodiversity-friendly landscapes. *Trends Ecol. Evol.* 36 (10), 919–930. <https://doi.org/10.1016/j.tree.2021.06.010>.
- Vehtari, A., Gelman, A., Gabry, J., Yao, Y., 2021. Package ‘loo.’ Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models.
- Verburg, P.H., Neumann, K., Nol, L., 2011. Challenges in using land use and land cover data for global change studies. *Glob. Chang. Biol.* 17 (2), 974–989. <https://doi.org/10.1111/j.1365-2486.2010.02307.x>.
- Waldner, F., Diakogiannis, F.I., 2020. Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network. *Remote Sens. Environ.* 245, 111741 <https://doi.org/10.1016/j.rse.2020.111741>.
- Waldner, F., Diakogiannis, F.I., Batchelor, K., Ciccotosto-Camp, M., Cooper-Williams, E., Herrmann, C., Mata, G., Toovey, A., 2021. Detect, consolidate, delineate: scalable mapping of field boundaries using satellite images. *Remote Sens. (Basel)* 13 (11). <https://doi.org/10.3390/rs13112197>.
- Wang, S., Waldner, F., Lobell, D.B., 2022. Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision. *Remote Sens. (Basel)* 14 (22). <https://doi.org/10.3390/rs14225738>.
- Weigel, R., Koellner, T., Poppenborg, P., Bogner, C., 2018. Crop diversity and stability of revenue on farms in Central Europe: an analysis of big data from a comprehensive agricultural census in Bavaria. *PLoS One* 13 (11), e0207454. <https://doi.org/10.1371/journal.pone.0207454>.
- Wesemeyer, M., Kamp, J., Schmitz, T., Müller, D., Lakes, T., 2023. Multi-objective spatial optimization to balance trade-offs between farmland bird diversity and potential agricultural net returns. *Agr. Ecosyst. Environ.* 345, 108316. <https://doi.org/10.1016/j.agee.2022.108316>.
- Whitcraft, A.K., Becker-Reshef, I., Justice, C.O., Gifford, L., Kavvada, A., Jarvis, I., 2019. No pixel left behind: toward integrating earth observations for agriculture into the United Nations sustainable development goals framework. *Remote Sens. Environ.* 235, 111470 <https://doi.org/10.1016/j.rse.2019.111470>.
- Williams, T.G., Bui, S., Conti, C., Debonne, N., Levers, C., Swart, R., Verburg, P.H., 2023. Synthesising the diversity of European Agri-food networks: A meta-study of actors and power-laden interactions. *Glob. Environ. Chang.* 83, 102746 <https://doi.org/10.1016/j.gloenvcha.2023.102746>.
- Wolz, A., 2013. The organisation of agricultural production in East Germany since World War II: Historical roots and present situation. Discussion Paper, No. 139. Leibniz Institute of Agricultural Development in central and Eastern Europe (IAMO), Halle (Saale).
- Wolz, A., Kopsidis, M., Reinsberg, K., 2010. The transformation of agricultural production cooperatives in East Germany and their future. *Journal of Rural Cooperation* 37.