A Flexible Approach for Quality Assurance in Geospatial Data: Mapping European Union Agricultural Census Data 2020*

Nicolas Lampach[a,c]†, Jon Olav Skøien[b], Helena Ramos[a], Rudolf Seljak[a],

Linda See[b], Renate Koeble[b] and Marijn Van Der Velde[b]

[a] *European Commission, Eurostat (E1), Luxembourg, Luxembourg*

[b] *European Commission, Joint Research Centre (D.5), Ispra, Italy*

[c] *University of Natural Resources and Life Sciences, Institute of Sustainable Economic Development, Vienna, Austria*

The combination of geographical information and official statistics is elementary to support evidence-based policies and improve effective monitoring in agriculture. Yet, public access to these fully coherent datasets at European Union (EU) level has been impeded owing to lack of appropriate methods assuring data quality. We develop a new flexible approach to rule out the disclosure of confidential and unreliable information from 9.1 million agricultural holdings when mapping the 2020 data of the EU agricultural census. Our approach leverages the power of multi-resolution grids by aggregating a risky cell at given resolution to the lowest resolution grid producing a non-confidential and reliable cell, while maximising the utility of the data. Our work discusses several aspects of the quality assurance framework ranging from coherence, consistency over time to accessibility of the new datasets. We demonstrate the approach by presenting a grid map on the share of organic farming for a selection of countries with confidentiality and reliability treatment. The implications of the method are discussed within the agricultural statistics literature and we provide an avenue for future research.

Keywords: Agricultural census; Maps; Multi-resolution; Organic farming; Quality; Spatial

*We thank the members of the working groups on farm statistics, innovation and methodology and geographical information system of the European Commission for their valuable feedback. Special thanks to our Danish colleagues for their support to attach a synthetic dataset of 2020 DK agricultural census as use-case to the R package *FSS*.

† *Corresponding author*: Nicolas Lampach, Eurostat, Unit E.1. Agricultural and Fisheries Statistics; E-mail: nicolas.lampach@ec.europa.eu.

1 Introduction

An agricultural census involves the regular and systematic collection of data on the structure of a nation's agricultural sector. The unit of data collection is the agricultural holding, which is comprised of the parcels of land and livestock managed by a single entity, such as an individual, household, or a public or private sector organization, for the purpose of agricultural production. By collecting information at regular intervals over time, such as the size of the holding, crop and livestock production and agricultural inputs, any changes in the agricultural sector can be monitored as well as their impacts on food security and the environment (FAO, 2017a).

In the European Union (EU), a decennial agricultural census is undertaken across Member States (MS) along with a sample survey every 3 to 4 years, referred to as the Farm Structure Survey[1] (FSS). Stipulated by Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018 on Integrated Farm Statistics (The European Commission, 2018), the collection of information in the FSS follows a common methodology in order to produce comparable and representative statistics across MS and over time. In the 2020 census, the FSS covered more than 9 million farm holdings.

FSS data are used to assess the state of agriculture across the EU, monitoring trends and transitions in the structure of farms (Eurostat, 2024). For example, Neuenfeldt et al. (2019) used FSS data to determine the drivers of farm structure change, finding that past farm structure explains the largest amount of variation but other drivers such as environmental conditions, prices, subsidies and income also play a role. The data are also key inputs to the management and evaluation of the Common Agricultural Policy (CAP) in terms of its environmental, economic and social impacts, and as inputs to CAP reforms, modelled, for example, using CAPRI (Barreiro-Hurle et al., 2021). In addition to the CAP, FSS data are valuable for other policy areas, including the environment, climate change, employment and regional development (e.g., Copus et al. (2006), Einarsson et al. (2020)).

The disclosure of micro-data from an agricultural census and survey data are additionally complicated, because agricultural holdings can contain information related to commercial operations or sensitive personal data. The release of census and survey data are, therefore, subject to confidentiality treatment, which states that data about individuals or enterprises cannot be released or disclosed. Statistical disclosure control is the process by which national statistical offices ensure that any confidentiality legislation is applied (Eurostat, 2019, FAO, 2017b). Different methods of statistical disclosure are used including table redesign such that information in the tables is aggregated, cell suppression in which values are completely omitted, and adjustment of values using different approaches such as rounding, controlled adjustment to replace cells with 'safe' values, and perturbation, where random noise is added to cells (European Commission, 2021, Fienberg and Jin, 2009, Hundepool et al., 2010, Quatember and Hausner, 2013, Templ, 2017).

Traditionally, the data is disclosed in tabular form by considering standard rules of confidentiality treatment

[1] The name originated from the former Regulation 1166/2008 on European farm survey (The European Commission, 2008) and most readers are familiar with this term. Current legislation (The European Commission, 2018) amended the name to Integrated Farm Statistics (IFS), which is less used, and therefore we opt to refer in this paper to FSS.

and quality rating system by aggregating to coarse administrative levels (i.e., NUTS2, NUTS1 or even national depending on the MS) before release by the European Statistical Office (Eurostat). However, there would be considerable value to policy design, policy impact assessment and scientific research more generally, in having access to data at a finer spatial resolution. Moreover, with advances in technology and the increasing trend to provide open access to government data across many sectors, new methods for disseminating data from censuses and surveys are needed (Shlomo, 2018).

To fill this gap, we present a methodology that applies legal confidentiality rules to individual farm records and produces non-confidential and reliable information on a multi-resolution spatial grid. We demonstrate the approach by displaying the share of organic area for limited number of countries. Although the methodology is specifically designed for FSS data, such an approach could also be easily adapted for releasing other individual census and survey based data that are subject to legal rules of disclosure.

2 Materials and methods

2.1 Farm structure survey data

European surveys on the structure of agricultural holdings have been carried out since 1966, and they aim to provide statistical knowledge for the monitoring and evaluation of related policies, in particular the CAP as well as environmental, climate change adaptation and land use policies. To reduce the burden on national administrations, Regulation (EU) 2018/1091 on integrated farm statistics provides a new framework by distinguishing between core and module variables[2], which vary in frequency and representativeness (The European Commission, 2018). It is required that the information on the core variables (e.g, general structural agricultural variables) should cover 98% of the utilised agricultural area and 98% of the livestock units of each MS. The modules contain information on specific topics such as the labour force, animal housing or irrigation, and can be carried out on samples of agricultural holdings by meeting the precision requirement laid down in Annex V of Regulation (EU) 2018/1091.

National data providers (i.e., national statistical institutes, ministries of agriculture or other governmental bodies) prepare the questionnaire, conduct the interviews and complete the survey with additional information from administrative registers (e.g wine, bovines, integrated information and the control system). The individual records at farm level are encrypted and transmitted to Eurostat via a secure system that implements an automated procedure to validate the content and structure of the micro data. For the first time during the 2020 agricultural census, Eurostat introduced an automated error detection procedure, leading to higher quality statistics.

The agricultural census for 2020 was carried out during the COVID-19 pandemic, posing significant challenges for countries to meet the official deadline of the data collection period. Although certain countries

[2]The complete list of variables surveyed during the European agricultural census 2020 can be found in the Implementing Regulation (EU) 2018/1874 of 29 November 2018 on the data to be provided for 2020 under Regulation (EU) 2018/1091 of the European Parliament and of the Council on integrated farm statistics and repealing Regulations (EC) No 1166/2008 and (EU) No 1337/2011, regarding the list of variables and their description.

experienced no impact from the pandemic due to early adoption of information and communication technologies (e.g, computer-assisted data collection mode, administrative registers), other countries were negatively impacted, particularly those that could not conduct face-to-face interviews due to social distancing restrictions. However, these exceptional circumstances also motivated the use of remote data collection methods (e.g., Computer assisted telephone/web interviewing, data collection by post) by one third of the countries.

Table 1: Data collection overview of the farm structure survey

Year	Type	Variables	Surveyed farms* (MM)	Population covered* (MM)	Countries
2010	Census	419	12.81	13.03	33
2013	Sample	358	1.73	11.04	30
2016	Sample	363	1.69	10.55	30
2020	Census	364	9.03	9.16	30

Note. *Covers all Member States, candidate and EFTA countries for the respective data collection year. Further details about the coverage can be found in Eurostat (2023a).

While an agricultural census is carried out every 10 years, sample surveys are administered during interim years. Table 1 summarises the data collection for the last decade by highlighting the number of variables, the number of surveyed farms, the population covered and the number of countries participating in the survey rounds. A substantial volume of information was collected during the 2020 survey campaign, which was comprised of more than 300 variables from around 9.03 million agricultural holdings. In sample survey years such as 2016, 1.69 million agricultural holdings were surveyed, which represents approximately 10.55 million holdings. It is worth mentioning that the accuracy and quality of the data is generally lower in the sample compared to the agricultural census. Therefore, we have undertaken a quality assessment of the indicators used in the production of the multi-resolution grid data; this will also ensure comparability.

2.2 Statistical disclosure and quality assurance framework

Official statistics are governed by a fundamental principle that protects the confidentiality of individuals or organisations and produces high-quality official statistics by masking sensitive information according to international and European law[3] (Eurostat, 2019, The European Commission, 2009, 2018, Trewin et al., 2007). At a higher spatial resolution, there is a legally binding obligation to employ appropriate aggregation and disclosure control to render spatial datasets accessible to the public (The European Commission, 2018). Furthermore, the Implementing Regulation (EU) 2018/1874 defines a set of rules for disclosing information from European surveys on the structure of agricultural holdings collected at farm location including the use of the 1 km INSPIRE Statistical Units Grid for pan-European data. In addition to the standard rules for tabular data, a key requirement is that values can only be disseminated at a 1 km grid when the cell includes more than ten agricultural holdings. Alternatively, aggregating to a nested 5 km or larger grid size is required to satisfy the aforementioned requirement (The European Commission, 2020).

[3] Separate national laws (EU/EEA/EFTA) might contain stricter (or laxer) rules related to the disclosure of personal information.

A disclosure occurs when an intruder correctly finds or determines some values about an individual or organization from the data released. Duncan and Lambert (1989) differentiate between two types of disclosure risk: identity disclosure and attribute disclosure. While the former occurs when a record can be directly linked to an individual, the latter refers to the knowledge gained about an individual or organization from the attribute(s) in the data released. Statistical Disclosure Control (SDC) techniques are widely deployed to reduce the risk of disclosing private information at an acceptable level, while maximising the utility of the data (Quatember and Hausner, 2013, Templ, 2017). From the two broad families of methods that exist, the perturbative method modifies the data prior to publication by adding random noise such as rounding to the nearest multiple of ten. Non-perturbative techniques reduce the amount of information by suppressing or aggregating the data. The optimal mixture of SDC should strike a balance between the mandatory privacy protection of the statistical output and the accessibility to the data at the highest available spatial resolution (Quatember and Hausner, 2013).

In agreement with member states, Eurostat (2020) has provided a series of recommendations in the confidentiality charter for disclosure control. For the dissemination of aggregated tabular statistics, key elements include mandatory compliance with the threshold[4] and dominance rule[5], and the statistical output must satisfy certain quality criteria[6]. These rules also apply to the dissemination of spatially explicit data on a grid.

Another important aspect that is receiving increasing attention is second order confidentiality, which occurs when the value of a suppressed sensitive cell can be determined from neighbouring cells or from other publicly available sources. In terms of gridded data, it is possible that cells become identifiable when both high-resolution gridded data and low-resolution NUTS data are published. Applying gap-filling methods to both datasets to impute the suppressed values would put the disclosure of private information at risk (Higgins and Scheiter, 2012). This threat can be overcome by carefully choosing the size of the grid cells and the type of administrative regions for the dissemination of the data.

Lastly, a quality control of the indicator is necessary. This is usually not an issue when (almost) the entire population has been sampled, as in census years, but in sample years the use of extrapolation factor (i.e weight) will introduce a prediction error to the produced indicator. The prediction error can be estimated as a function of the sample size, population size, sampled values and possible stratification. The Integrated Farm Statistics Manual (Eurostat, 2023b, Section 4.6) provides some guideline on the relative standard error (coefficient of variation) of the estimate that should be less than 0.35, otherwise the information is considered to be unreliable and thus cannot be disclosed.

2.3 Multi-grid resolution

Several different methods can be used to create a gridded data set that respect the quality assurance criteria, while all coming with certain advantages and disadvantages.

[4] Suppression of cells representing less than four agricultural holdings.

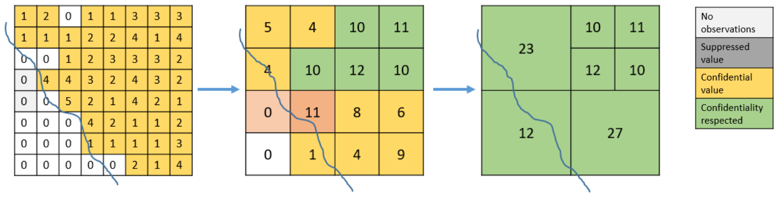
[5] Suppression of cells when one or two contributors are dominant.

[6] Data accuracy is evaluated based on sampling errors that can be estimated from the sample itself using the standard errors of the estimated values. If the coefficient of variation of the estimated values is larger than 35%, the cell is usually suppressed.

A less common approach constitutes the multi-resolution grid or quadtree whereas the grid size is a function of the variable of interest and follows a hierarchical data structure (Asim et al., 2023, Behnisch et al., 2013, Eurostat, 2020, Lagonigro et al., 2020). The main idea is that the resolution of the grid will vary according to the density of the observations to ensure that the confidentiality rules are respected for all grid cells. An example of this is shown in Figure 1 with the same fictitious data and the same aggregation (middle panel) as in the value suppression approach. However, when reducing the resolution towards the third and final map in Figure 1, the four cells in the upper right corner are not aggregated, as they already respect the confidentiality rules. Hence it is possible to share the data with a higher resolution in this area and at a coarser resolution in the rest of the map.

The method is sensitive to islands and borders, where it might be difficult to include a sufficient number of holdings, even for relatively large grid cells, when most of the grid cell does not include the data. Therefore, it might not be recommended to continue the process until absolutely all grid cells respect the confidentiality rules, but for some resolution to suppress the values from grid cells that still do not respect the confidentiality rules. Similar to the value suppression method, this means that the sum of the values from the grid cells will be lower than the total number of holdings. However, this need for suppression will affect considerably fewer grid cells than when using suppression on a regular grid. This method is, therefore, the one that has been chosen for the high resolution gridded data from the FSS.

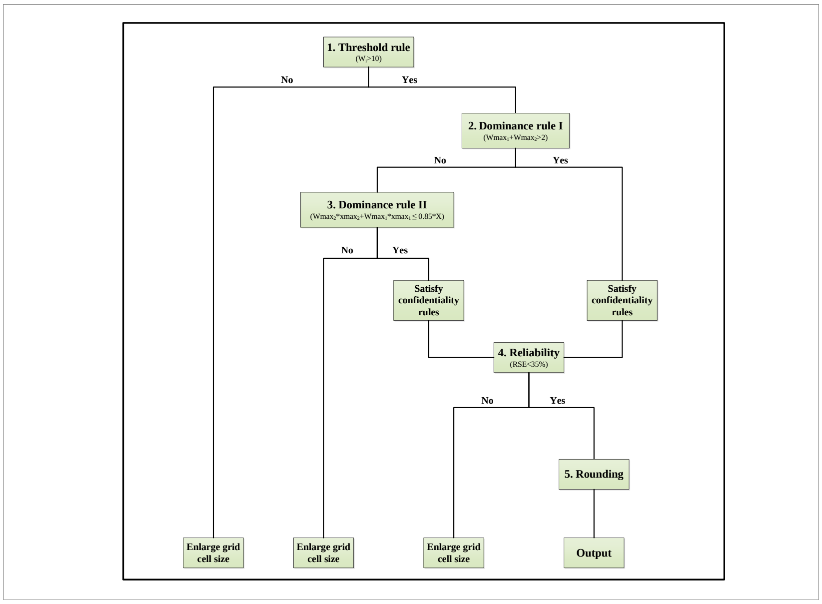
Figure 1: Example of a multi-resolution grid, moving from a higher to a coarser resolution.



Note. The numbers represent the number of holdings per grid cell. The line shows a border effect (regional border or land/sea).

There are some cases where aggregation might not be desired. In the situation where a relatively large single grid cell does not respect the confidentiality rules, it is fine to aggregate it if the neighbouring grid cells are also relatively large. However, it would be unfortunate if the single cell was aggregated with many smaller grid cells that could otherwise be disseminated at a high resolution. The added value of being able to present a value for a region with very few farms is perhaps lower than what is lost by having to aggregate to a lower resolution. We have, therefore, also introduced the possibility for the user to set a minimum share of a grid cell value relative to the possible lower resolution grid cell before it is necessary to aggregate. The check will either be against the variable of interest (utilised agricultural area, livestock units, etc.) or the number of holdings if no variable is analysed. If the limit is 0.1 (although a smaller value is recommended), the grid cell with 1 in the lower left quadrant would not lead to aggregation in Figure 1, as it represents less than 10% of the value of the possible lower resolution grid cell. Instead, it would be left and suppressed in the later step.

Figure 2: Flowchart showing the rules that are applied for the release of FSS data.



Note. Where the rules are not satisfied, the grid cell sizes must be increased as implemented in the multi-grid solution.

Figure 2 displays the iterative process of producing a nested structure of multi-hierarchical grids satisfying a set of confidentiality rules and quality requirements. We denote the level of resolution $k \in K$ with $K = \{k_0, k_1, \ldots, k_m\}$ where k_0 is the highest resolution (1 km for FSS) and k_m the lowest resolution. The iteration starts with i_1, the possible aggregation from k_0 to k_1 and continues until reaching the maximum level k_m $(i \in \{i_1, \ldots, i_m\})$.

The confidentiality rules are evaluated in the following order, where the threshold rule must be passed, whereas it is sufficient to pass one of the dominance rules:

1. Threshold rule:

   If the aggregated extrapolated number of agricultural holdings in grid cell l (w_l) for resolution k_0 in iteration i_1 is less than ten ($W_l < 10$ with $W_l = \sum_{j=1}^{n_l} w_j$) where n_l is the number of records in l, then the grid cell size must be enlarged to k_1 and the confidentiality rules for the new grid cell will be assessed in iteration i_2;

2. Dominance rule I:

   If, after ordering the variable of interest in descending order, the sum of the weights w_{jmax_1} of the highest value x_{max_1} ($W_{max_1} = w_{jmax_1} \times x_{max_1}$) and of the second highest value x_{max_2} ($W_{max_2} = w_{jmax_2} \times x_{max_2}$) is greater than two ($W_{max_1} + W_{max_2} > 2$), then the dominance rules are satisfied for the grid cell at k_0 and the reliability of the results needs to be assessed. (Note that the weights are rounded before this step, so larger than 2 means at least 3). Otherwise Dominance rule II needs to be satisfied;

3. Dominance rule II:

   If the two potential dominant contributors are less than or equal to 85% of the extrapolated aggregated value of the grid cell ($W_{max_2} \times x_{max_2} + W_{max_1} \times x_{max_1} \leq 0.85 \times X$), then the confidentiality rules are satisfied; otherwise the grid cell size must be enlarged to k_1 and the confidentiality rules for the new grid cell will be assessed in iteration i_2;

4. Reliability of the results:

   If the coefficient of variation (Relative Standard Error (RSE)) for the grid cell at k_0 is less than 35%, then the indicator is reliable (to be disseminated with a warning if above 25%); otherwise the grid cell size must be enlarged to k_1 and and the confidentiality rules for the new grid cell will be assessed in iteration i_2;

5. Rounding:

   After the last iteration, and as a measure to add further perturbation to the disclosed information, all non-confidential extrapolated number of holdings and extrapolated aggregated values of variables are rounded to the nearest multiple of ten.

3 Results

To demonstrate our approach, we employ the share of organic farming in utilised agricultural area -crucial indicator to monitor the CAP and objective of European Green Deal- by producing jointly the grid cells for both variables of interest, organic area under conversion and fully converted to organic farming and utilised agricultural area and subsequently calculating the ratio for each grid cell.

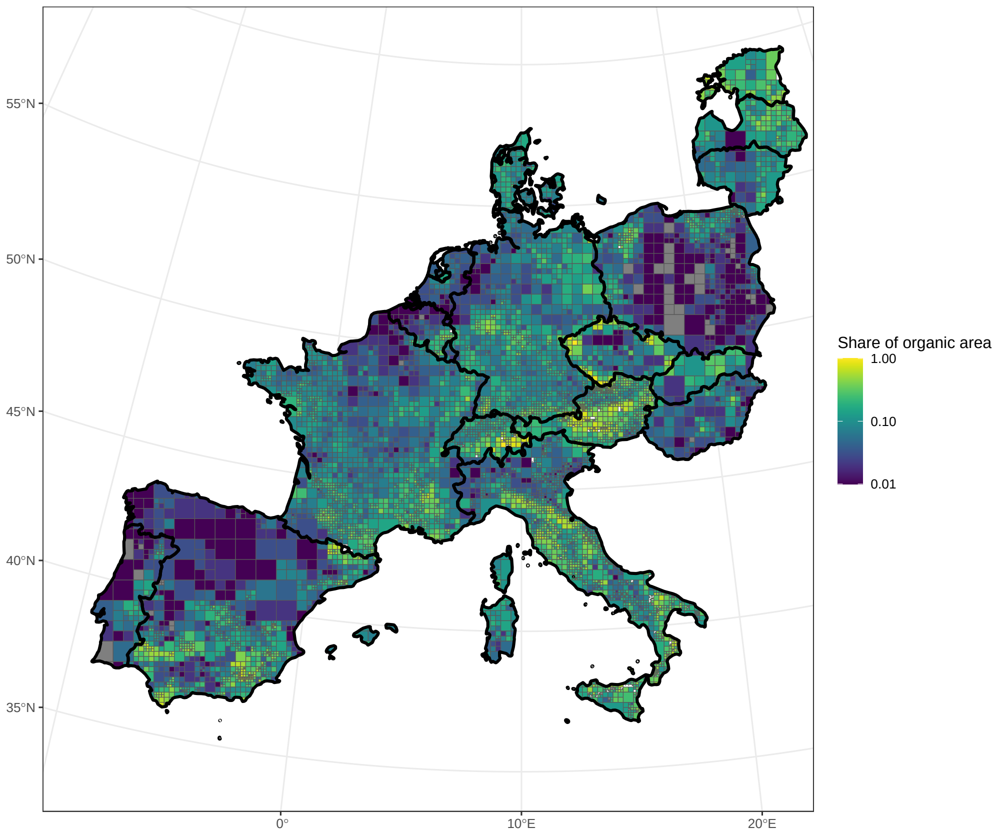


Figure 3: Share of organic area in utilised agricultural area for a selection of countries in 2020

Note. Grid cells representing less than 5% of the aggregated value of lower resolution are suppressed (greyish cells). Similar those grid cell that reach the maximum level of resolution without complying with statistical disclosure rules are also suppressed.

Displayed in Figure 3 is the share of organic farming in 2020 for a selection of countries. We observe that the size of grid cell is a function of the density of organic farms and the color graduation captures the magnitude of the variable of interest. Furthermore, the combined approach handles relatively well border and coastline effects.

Obviously, there are spatial structural differences within and across countries. For instance, there is a greater density of organic farms and a higher usage of organic farming located in southern part of France, Spain and partly in Germany than in northern location of the respective country. Moreover, Poland and Hungary display a heterogenous pattern with a lower degree in the usage of organic farming in comparison to Estonia and Latvia joining the European Union at the same moment.

4 Conclusion

The multi-resolution gridded solution presented here represents a step change in the way that the rich amount of information on the farming sector in Europe, collected by EU Member States and Eurostat in agricultural censuses and surveys, could be released in the future. The method aims at producing gridded layer of varying resolutions that maximizes the information content at an aggregated level while assuring quality data. In contrast, other countries outside of the EU are still much stricter in their dissemination of agricultural census data. For example, the United States Department of Agriculture releases data at county level, which is similar to NUTS2 regions in Europe USDA NASS (2024). In Canada, one-third of data were not disclosed in the 2016 agricultural census, which employed suppression of data. For the 2021 Census, Statistics Canada has switched to the use of random tabular adjustment, which makes changes to individual cells to ensure data protection Statistics Canada (2023). Hence, the suggested approach could be used and adapted by other statistical services that disseminate agricultural census and survey data (such as farm accountancy data) to meet their specific disclosure requirements. Given the versatile and flexible implementation of our approach, the methodology could easily be expanded to other statistical domains where sensitive information on individuals or enterprises is collected, such as population, migration, business and labour force statistics.

The next steps are to apply the method to produce a set of key agricultural indicators from the agricultural census and survey data for Europe, which can be used to better understand agricultural systems across Europe and to identify the main drivers of the adoption of agricultural practices. The release of grids for analyzing changes over time will be more challenging as the multi-resolution grids will need to be spatially consistent if meaningful comparisons are to be made. Methods for ensuring both spatial and temporal consistency will be added in the future.

References

Asim, K., Schorlemmer, D., Hainzl, S., Iturrieta, P., Savran, W., Bayona, J., and Werner, M. (2023). Multi-Resolution Grids in Earthquake Forecasting: The Quadtree Approach. *Bulletin of the Seismological Society of America*, 113(1):333–347.

Barreiro-Hurle, J., Bogonos, M., Himics, M., Hristov, J., Pérez-Domínguez, I., Sahoo, A., Salputra, G., Weiss, F., Baldoni, E., and Elleby, C. (2021). Modelling environmental and climatic ambition in the agricultural sector with the CAPRI model: exploring the potential effects of selected farm to fork and biodiversity strategies targets in the framework of the 2030 climate targets and the post 2020 Common Agricultural Policy. JRC Technical Report, Publications Office, Luxembourg.

Behnisch, M., Meinel, G., Tramsen, S., and Diesselmann, M. (2013). Using quadtree representations in building stock visualization and analysis. *Erdkunde*, 67(2):151–166.

Copus, A., Hall, C., Barnes, A., Dalton, G., Cook, P., Weingarten, P., Baum, S., Stange, H., Lindner, C., Hill, A., Eiden, G., McQuaid, R., Grieg, M., and Johansson, M. (2006). Study on Employment in Rural Areas. Available at: https://www.napier.ac.uk/˜/media/worktribe/output-246104/serareport1ruraleu272006pdf.pdf. Technical report.

Duncan, G. and Lambert, D. (1989). The risk of disclosure for microdata. *Journal of Business & Economic Statistics*, 7(2):207–217.

Einarsson, R., Pitulia, D., and Cederberg, C. (2020). Subnational nutrient budgets to monitor environmental risks in EU agriculture: calculating phosphorus budgets for 243 EU28 regions using public data. *Nutrient Cycling in Agroecosystems*, 117(2):199–213.

European Commission (2021). Statistical disclosure control. In *European Business Statistics Manual: 2021 Edition*. Publications Office, LU.

Eurostat (2019). Quality assurance framework of the european statistical system. Available at: https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V2.0-final.pdf.

Eurostat (2020). Integrated Farm statistics manual - 2020 edition. Technical report, Eurostat.

Eurostat (2023a). Glossary:farm structure survey (fss). Available at: urlhttps://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Farm$_structure_survey_(FSS)$.

Eurostat (2023b). Integrated Farm statistics manual - 2023 edition. Technical report, Eurostat.

Eurostat (2024). Farm Structure Survey. Available at: https://ec.europa.eu/eurostat/web/microdata/farm-structure-survey.

FAO (2017a). *World programme for the census of agriculture 2020. Volume 1 Programme, concepts and definitions*. Number 15 in FAO Statistical Development Series. Food and Agriculture Organization of the United Nations, Rome. OCLC: 1091006801.

FAO (2017b). *World programme for the census of agriculture 2020. Volume 2 Operational guidelines*. Number 16 in FAO Statistical Development Series. Food and Agriculture Organization of the United Nations, Rome. OCLC: 1091006801.

Fienberg, S. E. and Jin, J. (2009). Statistical Disclosure Limitation For Data Access. In Liu, L. and Özsu, M. T., editors, *Encyclopedia of Database Systems*, pages 2783–2789. Springer US, Boston, MA.

Higgins, S. I. and Scheiter, S. (2012). Atmospheric CO2 forces abrupt vegetation shifts locally, but not globally. *Nature*, 488(7410):209–212.

Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Lenz, R., Longhurst, J., Nordholt, E. S., Seri, G., and Wolf, P. (2010). Handbook on statistical disclosure control. *ESSnet on Statistical Disclosure Control*.

Lagonigro, R., Oller, R., and Martori, J. C. (2020). AQuadtree: an R Package for Quadtree Anonymization of Point Data. *The R Journal*, 12(2):209–225.

Neuenfeldt, S., Gocht, A., Heckelei, T., and Ciaian, P. (2019). Explaining farm structural change in the European agriculture: a novel analytical framework. *European Review of Agricultural Economics*, 46(5):713–768.

Quatember, A. and Hausner, M. C. (2013). A family of methods for statistical disclosure control. *Journal of Applied Statistics*, 40(2):337–346.

Shlomo, N. (2018). Statistical disclosure limitation: New directions and challenges. *Journal of Privacy and Confidentiality*, 8(1).

Statistics Canada (2023). Guide to the Census of Agriculture, 2021. Available at: https://www150.statcan.gc.ca/n1/pub/32-26-0002/322600022021001-eng.htm.

Templ, M. (2017). Statistical disclosure control for microdata. *Cham: Springer*.

The European Commission (2008). REGULATION (EU) 2008/1166 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 19 November 2008 on farm structure surveys and the survey on agricultural production methods and repealing council regulation (EEC) No 571/88. *Official Journal L321*, 03:14–34.

The European Commission (2009). Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics and repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the transmission of data subject to stat. *Official Journal L87*, 52:164–173.

The European Commission (2018). REGULATION (EU) 2018/1091 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 July 2018 on integrated farm statistics and repealing Regulations (EC) No 1166/2008 and (EU) No 1337/2011. *Official Journal L200*, 61:1–29.

The European Commission (2020). COMMISSION IMPLEMENTING REGULATION (EU) 2018/1874 of 29 November 2018 on the data to be provided for 2020 under Regulation (EU) 2018/1091 of the European Parliament and of the Council on integrated farm statistics and repealing Regulations (EC) No 1166/20. *Official JOurnal L306*, Novermber 2018:14–19.

Trewin, D. et al. (2007). Principles and guidelines of good practice for managing statistical confidentiality and microdata access. *UNECE United Nations Economic commission for Europe: http://www. unece. org/stats/documents/tfcm/1. e. pdf*.

USDA NASS (2024). Census of Agriculture. Available at: https://www.nass.usda.gov/{Publications}/{AgCensus}/2022/index.php.