D2.2 Publication of farm typologies across the EU

Preliminary results and findings

Dissemination Level	☒ Public ☐ Sensitive
Version	V1.1 Approved by the EU

Project Acronym	Europe-LAND
Work Package	WP2 Revealing agricultural land-use behaviour and its drivers in Europe
Lead Author (Partner)	Kristoffer Ansbak Petersen (04/UCPH), Martin Rudbeck Jepsen (04/UCPH)
Deliverable	Publication of farm typologies across the EU
Deliverable Lead	04/UCPH
Type	Data – data sets, microdata, etc
Contractual delivery date	31.05.2026
Actual submission date	29.05.2026

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Revision and history chart

Version	Date	Main author	Summary of changes
0.1	21.05.2026	Kristoffer Ansbak Petersen, Martin Rudbeck Jepsen	Draft created
0.2	27.05.2026	Lucie Kupková	Deliverable checkpoint
0.3	28.05.2026	Kristoffer Ansbak Petersen	Minor text edits
0.4	28.05.2026	Franziska Wolf	Peer review, comments
0.5	28.05.2026	Kristoffer Ansbak Petersen	Minor text edits
0.6	29.05.2026	Minami Tsuchikura	Final quality control
1.0	29.05.2026	Minami Tsuchikura	Submit to the EU portal
1.1	20.10.2026	Minami Tsuchikura	Final approval by EU received

Continuation of the Work

This report was prepared as part of Deliverable D2.2 (Publication of farm typologies across the EU) of the Europe-LAND project. The report contains preliminary and unpublished draft material intended for reporting and demonstration purposes only. The work will be further continued and developed, and the methods, results, and conclusions presented here may therefore be revised, updated, or modified prior to subsequent publication in a peer-reviewed scientific journal. The findings presented should therefore not be interpreted as definitive at this stage.

Limitations

The study does not cover the entire EU but includes data from 18 EU Member States (Austria, Belgium, Cyprus, Czechia, Denmark, Estonia, Germany, Ireland, Italy, Latvia, the Netherlands, Poland, Portugal, Romania, Slovenia, Slovakia, Spain, and Sweden) and therefore provides broad coverage of European farming systems. Furthermore, Belgium, Germany, and Italy are only partially represented due to limited data availability at the regional or federal state level. Consequently, the resulting farm typologies should not be interpreted as a complete representation of agricultural systems across the entire EU.

Disclaimer

This document does not represent the opinion of the European Union or EC-CINEA, and the European Community or EC-CINEA is not responsible for any use that might be made of its content. This document may contain material that is the copyright of certain Europe-LAND consortium parties and may not be reproduced or copied without permission. All Europe-LAND consortium parties have

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

agreed to the full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the Europe-LAND consortium nor a certain party of the Europe-LAND consortium warrant that the information contained in this document is capable of use nor that use of the information is free from risk and does not accept any liability for loss or damage suffered by any person using this information.

Acknowledgement

This project has received funding from the European Commission's Horizon Europe Research and Innovation Programme under grant agreement No. 101081307.

Further, we would like to explicitly acknowledge the valuable support of all our consortium partners in reaching out to their national, regional and local authorities to obtain IACS data. Finally, we appreciate very much the support of and thank all the data providing entities in EU member states that shared IACS datasets with us.

Table of content

List of Abbreviations and Acronyms 5
1. Introduction 6
2. Data and methodology 6
2.1. Data 6
2.2. Farm Typology Construction 9
3. Results 11
3.1. Overview of farm typologies 11
3.2. Structural characteristics of farm typologies 13
3.3. Spatial distribution of farm typologies 14
4. Concluding remarks 15
5. References 16

List of Tables

Table 1: Overview of the crop groups used for farm typology construction and principles for aggregation applied to the harmonised HCAT classifications. ..... **Error! Bookmark not defined.**

Table 2: Overview of country-level data coverage, number of farms, and agricultural area used for typology construction. ..... **Error! Bookmark not defined.**

List of Figures

Figure 1:Elbow plot of WCSS against the number of clusters (k). The optimal number of clusters was determined using the Kneedle algorithm, which identifies the point of maximum curvature in the curve. This point indicates where additional clusters provide diminishing improvements in cluster separation. The analysis identified k = 16 as the selected number of clusters. 11

Figure 2: Relative crop composition portfolios of the 16 farm typologies identified using two-stage clustering of harmonised GSA data. Farm typologies broadly represent extensive production systems (Clusters 1 and 11), mixed grassland-cereal and fodder-oriented systems (Clusters 2, 3, 5, and 8), cereal- and maize-dominated systems (Clusters 4, 6, and 7), specialised crop systems (Clusters 10, 13, 14, and 15), and diversified mixed farming systems (Clusters 9, 12, and 16). 12

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Figure 3: Structural characteristics of the identified farm typologies. Bubbles represent the dominant crop groups in each typology, while bubble size corresponds to the number of farms within each typology. 13

Figure 4: Spatial distribution of the 16 farm typologies derived from harmonised GSA data across 18 EU Member States. 14

List of Abbreviations and Acronyms

CAP	Common Agricultural Policy
EU	European Union
GSA	GeoSpatial Application
HAC	Hierarchical agglomerative clustering
HCAT	Hierarchical Crop and Agriculture Taxonomy
IACS	Integrated Administration and Control System
SOM	Self-organizing maps
WCSS	Within-cluster sum of squares

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

1. Introduction

Detailed and harmonised information on agricultural land use is essential for understanding structural differences in European farming systems and for supporting agricultural, environmental, and climate policy. Within the European Union (EU), the Integrated Administration and Control System (IACS) provides a major administrative framework for the collection and management of agricultural data. Its associated GeoSpatial Application (GSA) datasets enable spatially explicit analyses by linking reported crop declarations to georeferenced agricultural parcels and farm identifiers. Compared with traditional statistical datasets, GSA data offer a higher level of spatial and thematic detail, enabling analyses of land use and production systems at the field and farm level (Jänicke et al., 2024; Leonhardt et al., 2024).

Despite this potential, the use of IACS, and particularly GSA data, in cross-country research remains limited. Because national GSA datasets are primarily developed for administrative purposes, they vary considerably across EU Member States in crop classification systems, reporting practices, and data accessibility (Jänicke et al., 2025; Leonhardt et al., 2024). These inconsistencies continue to pose challenges for data interoperability and cross-country comparisons, reducing the feasibility of EU-wide agricultural analyses.

Recent harmonisation initiatives, including the Europe-LAND and EuroCrops projects, have sought to address these challenges by translating national crop codes of GSA data into more consistent classification systems (Jänicke et al., 2025; Schneider et al., 2023). These datasets create new opportunities for cross-country agricultural analysis within the EU, allowing analyses to move beyond national statistics toward more detailed assessments of farm structures, land-use patterns, and production specialisation. However, harmonisation alone cannot fully eliminate the underlying differences in reporting practices, classification systems, and administrative objectives that exist across national datasets. These limitations highlight the continued need for more standardised approaches to agricultural data reporting and classification across EU Member States.

This work builds on these developments by assessing the feasibility of constructing farm typologies from harmonised GSA datasets covering 18 EU Member States. Field-level crop information was aggregated to the farm level in order to characterise farms according to their crop composition portfolios. The availability of farm identifiers for all reported parcels provides a unique opportunity to analyse not only how individual parcels are managed, but also which fields are managed by the same farm holding. This enables insights into the crop composition portfolios and management choices of millions of farm holdings across Europe.

Given the differences in national crop classification systems, the study adopts a pragmatic crop aggregation approach designed to improve cross-country comparability while explicitly acknowledging classification-related uncertainties. Rather than aiming for fully standardised farm classifications, the analysis focuses on evaluating whether harmonised farm-level datasets can support agricultural analysis across EU Member States.

2. Data and methodology

2.1. Data

This study uses harmonised GSA data from the internal database of the Europe-LAND project (Jänicke et al., 2026). The dataset contains field-level agricultural land-use information, including reported crop types for individual parcels, together with unique farm identifiers that enable aggregation of

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

parcels to the farm level. Only records containing a farm identifier were included in the analysis, as farm typologies were constructed at the farm level. Data from 18 EU Member States were included, providing broad coverage of European farming systems. Of the 18 countries, Belgium, Germany, and Italy only include partial regional coverage due to limited data availability at the regional or federal state level. The harmonised data build upon the Hierarchical Crop and Agriculture Taxonomy (HCAT) developed within the EuroCrops project, which translates national crop classifications into a hierarchical and machine-readable crop classification system (Schneider et al., 2023). Several refinements and extensions were implemented within the taxonomy to improve thematic resolution and consistency across land-use and crop-type categories for the purpose of farm typology construction (Jänicke et al., 2026).

Harmonised crop classifications were aggregated into broader crop groups in order to improve comparability across national datasets (Table 1). Parcel areas were subsequently used to calculate the proportional distribution of crop groups within individual farms, and each farm was assigned a geolocation based on the centroid of all parcels sharing the same farm identifier. Crop classes assigned to the not_included category were excluded from the modelling, as they primarily represent non-agricultural land, non-productive areas, or residual classes not directly relevant to farm typologies.

Data from the Common Agricultural Policy (CAP) 2023–2027 programme period were used whenever available. Where these datasets were not accessible, the most recent available GSA data were used instead (Table 2). Thus, the dataset spans multiple reporting years (2022–2025), which may introduce minor temporal inconsistencies due to differences in CAP implementation, reporting practices, or annual agricultural conditions. However, as the analysis focuses on overall farm crop composition portfolios and structural and spatial patterns rather than year-specific dynamics, the influence of these temporal differences is expected to be limited. The final dataset used for modelling after excluding records without farm identifiers and farms consisting exclusively of not_included land categories comprises more than 3.6 million individual farms covering almost 90 million hectares of European farmland.

Despite the harmonisation efforts, several data limitations remain. National IACS datasets differ in their crop classification systems in terms of classification detail and thematic scope. Crop categories that appear similar across countries may not necessarily represent the same agricultural function, as they may reflect varying combinations of land cover, land use, and administrative reporting objectives, making consistent harmonisation difficult. In addition, certain crop categories may be absent, aggregated differently, or reported inconsistently across countries, limiting full comparability.

Additional uncertainty may arise from the translation and harmonisation of national crop classes into the HCAT taxonomy and subsequent crop groups. Due to the diversity and ambiguity of national crop descriptions, some crops may be translated incorrectly or assigned to crop categories that do not fully reflect the characteristics of the original reported categories. As a result, some harmonised crop categories may differ in their underlying agricultural meaning across countries, which should be considered when interpreting the resulting typologies and spatial patterns.

Despite these limitations, the harmonised datasets still provide a basis for analysis of farm-level land-use patterns across EU Member States. The overall analytical workflow used to construct farm typologies from harmonised GSA data consists of four main steps: (1) grouping data into broader crop groups, (2) constructing farm-level crop composition portfolios, (3) clustering farm portfolios, and (4) deriving and interpreting farm typologies and their spatial distributions.

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Table 1: Overview of the crop groups used for farm typology construction and principles for aggregation applied to the harmonised HCAT classifications.

Crop group	**Principles for aggregation**
Cereals	Annual grasses grown for dry grain (e.g. wheat, barley, oats, rice, rye, triticale, millet, sorghum, pseudocereals), including mixed or undersown systems where grain is the primary product, excluding maize.
Maize (grain)	Maize grown for dry grain, excluding silage maize.
Protein crops	Annual legumes grown for dry seeds, including beans, peas, lentils, chickpeas, lupins, and soybeans.
Oilseeds	Crops grown for seed oil production, including rapeseed, sunflower, poppy, linseed, mustard, and camelina.
Industrial crops	Crops grown for industrial, energy, fibre, medicinal, aromatic, or biochemical uses (e.g. hemp, tobacco, herbs, spices, mushrooms).
Root crops	Crops grown for roots or tubers, including potatoes, sugar beet, fodder beet, and root vegetables.
Plants harvested green	Crops harvested green for fodder or biomass production, including temporary grass, forage legumes, green-cut cereals, silage maize, and crop mixtures.
Permanent grassland	Long-term grassland outside regular crop rotations, primarily used for grazing or mowing.
Vegetables and horticulture	Intensively managed crops for fresh consumption or ornamental use, including vegetables, herbs, flowers, nurseries, and greenhouse production.
Permanent crops	Long-term perennial crops producing fruits, nuts, or berries, including orchards, vineyards, and olive groves.
Woody vegetation	Woody vegetation not classified as permanent crops, including forests, agroforestry systems, and short-rotation coppice.
Fallow	Arable land temporarily left uncultivated within crop rotations, including set-aside land.
Not included	Non-agricultural land, non-productive areas, unknown categories, and residual classes (e.g. infrastructure, landscape features, wetlands).

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Table 2: Overview of country-level data coverage, number of farms, and agricultural area used for typology construction.

Country	2022	2023	2024	2025	Farms	Area [ha]
Austria	✓				110,334	3,170,245
Belgium*			(✓)		14,155	747,101
Cyprus			✓		28,344	126,561
Czechia				✓	33,656	3,504,209
Germany*	(✓)	(✓)		(✓)	215,591	12,200,224
Denmark				✓	30,809	2,507,318
Estonia			✓		7,160	899,857
Spain			✓		630,508	23,278,897
Ireland			✓		123,685	5,038,608
Italy*		(✓)	(✓)		94,925	2,164,332
Latvia		✓			51,088	1,761,132
Netherlands			✓		48,494	1,796,815
Poland			✓		1,205,599	13,971,652
Portugal				✓	171,559	3,417,822
Romania		✓			736,542	9,903,195
Sweden		✓			54,800	2,924,336
Slovenia			✓		53,354	461,979
Slovakia				✓	13,679	1,768,395
Total	2	5	9	5	3,624,282	89,642,678

✓ = data available for the whole country; (✓) = data available for selected regions only (see details below).

For Belgium, data are available from Wallonia (2024). For Germany, data are available from eight federal states: BAV (2025); BRB (2022); and LSA, MWP, NRW, RLP, SAT, and THU (2023). For Italy, data are available from three regions: EMR (2024); and MAR and TOS (2023).

2.2. Farm Typology Construction

The first step in the farm typology construction was to aggregate the field-level crop information to the farm level using the farm identifiers provided in the harmonised GSA datasets. For each farm, the area assigned to each crop group was summed across all parcels, resulting in a vector of farm-level crop composition portfolio describing the distribution of agricultural area among crop groups. Because individual farms cultivate only a limited number of crops, whereas the crop composition portfolio includes all crop groups represented in the harmonised dataset, the resulting farm vectors contain a large number of zero values.

To compare crop composition portfolios independently of total farm size, data were transformed using chord normalisation (Legendre & Gallagher, 2001). With chord normalisation, each farm's vector of crop areas is scaled to length 1, such that Euclidean distances between transformed vectors correspond to chord distances between the original farm-level crop composition portfolios. The transformation preserves zero values, emphasises relative composition rather than absolute area, and is compatible with Euclidean distance-based multivariate methods such as clustering. Chord

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

normalisation is commonly recommended for ecological abundance data (Legendre & Gallagher, 2001), which, like crop composition portfolios, are compositional in nature and characterised by non-negative values, a high frequency of zeros, and uneven category distributions. It is therefore well suited for exploratory analysis of farm-level crop composition portfolios, although it does not constitute a formal compositional data analysis approach. Formal log-ratio transformations were not adopted because of the high prevalence of structural zeros and the potential distortions introduced through zero replacement.

Farm typology construction followed a two-stage clustering approach consisting of (1) mapping farms onto representative prototype vectors using self-organizing maps (SOMs) (Kohonen, 2001), followed by (2) hierarchical agglomerative clustering (HAC) of SOM codebook vectors to derive farm typologies. Compared with direct clustering of the original data, this two-stage approach has been shown to perform well while reducing computational requirements (Vesanto & Alhoniemi, 2000), and it has also been applied in other archetype clustering studies (Beckmann et al., 2022). The SOM is an unsupervised neural network approach that projects high-dimensional input data onto a two-dimensional grid while preserving similarity relationships between observations (Kohonen, 2001). It was trained using the chord-transformed crop composition portfolios, allowing farms with similar crop composition portfolios to be positioned close to one another on the map.

The SOM grid size and structure were determined using commonly applied heuristic approaches based on the number of observations and the variance structure of the input data (Vesanto & Alhoniemi, 2000). The number of map units was estimated using the equation $m = 5 \sqrt{n}$, where n represents the number of farms included in the analysis. Grid dimensions were then determined according to the ratio between the two largest eigenvalues of the input data. Following this approach, a hexagonal SOM grid with dimensions of 112 x 85 was selected, corresponding to 9,520 map units.

The trained SOM was subsequently grouped into typology classes by clustering the SOM prototype vectors using Ward's HAC method (Ward, 1963). To determine the number of typology classes, the within-cluster sum of squares (WCSS) was computed for clusters ranging from $k = 1$ to $k = 100$. The optimal number of clusters was identified using the Kneedle algorithm for automated elbow-point detection (Satopaa et al., 2011), which detected an elbow at $k = 16$ (Figure 1). The SOM units were therefore grouped into 16 typology classes, and farms were assigned to typology classes according to the cluster membership of their best-matching SOM unit.

Typology-specific crop composition portfolios and descriptive summaries were subsequently constructed at the type-level and linked to all farms belonging to the corresponding type. For each farm typology, the dominant crop group, mean and variance of farm size, number of fields, and number of farms were calculated. These summaries were used to characterise differences in farm structure across typologies and to support the interpretation of the mapped spatial patterns. To analyse the spatial distribution of farm typologies, farm-level data were aggregated to a 1 km hexagonal grid covering the EU, with each hexagon assigned the area-dominant typology.

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

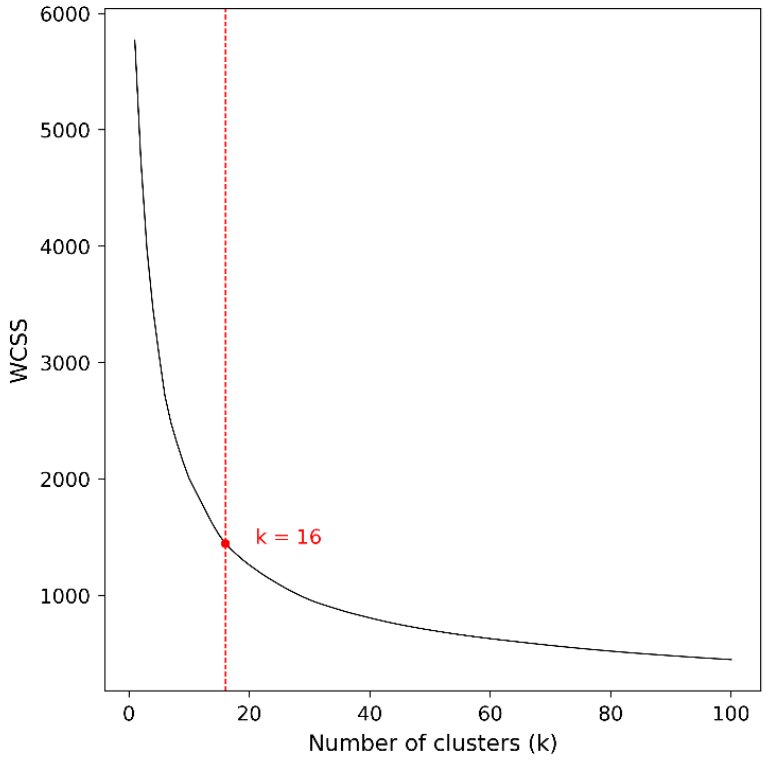


Figure 1:Elbow plot of WCSS against the number of clusters (k). The optimal number of clusters was determined using the Kneedle algorithm, which identifies the point of maximum curvature in the curve. This point indicates where additional clusters provide diminishing improvements in cluster separation. The analysis identified k = 16 as the selected number of clusters.

3. Results

3.1. Overview of farm typologies

The two-stage clustering approach identified 16 farm typologies characterised by distinct crop composition portfolios (Figure 2). The typologies capture a broad range of farm behavioural patterns, ranging from highly specialised farms dominated by a single crop group to more diversified production systems. They also reflect variation in production intensity, from extensive, low-intensity farming systems to more intensive crop-dominated systems. Intensity levels are inferred based on expected input and output intensities (Erb et al., 2013; Kuemmerle et al., 2013).

Extensive production systems are primarily represented by the permanent grassland cluster (Cluster 1), characterised by large shares of permanent grassland and limited crop diversification. Other grassland-dominated clusters (Clusters 2, 5, and 8) combine permanent grassland with cereals, maize, or fodder crops, reflecting more diversified and moderately intensive production patterns, likely related to livestock production. Another relatively extensive system is represented by Cluster 11, which is characterised by large shares of fallow land, indicating lower-intensity land use.

More intensive and specialised systems are observed among cereal-dominated farms (Clusters 6–7), with Cluster 7 showing particularly high specialisation and limited diversification. Maize-oriented farms (Cluster 4) are similarly characterised by large shares of grain maize and relatively intensive arable production. Cluster 3 is characterised by plants harvested green and likely reflects fodder-

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

oriented systems associated with livestock production. Additional specialised typologies include woody vegetation systems (Cluster 10), permanent crop farms (Cluster 13), vegetable and horticultural farms (Cluster 14), and industrial crop systems (Cluster 15). These typologies reflect varying levels of production intensity, ranging from relatively extensive woody vegetation systems to highly intensive vegetable and horticultural production.

Mixed farming systems are represented by Clusters 9, 12, and 16, which combine cereals with either protein crops, root crops, or oilseeds. These typologies follow traditional crop rotation (Jepsen et al., 2015) and reflect more diversified production structures with intermediate levels of production intensity.

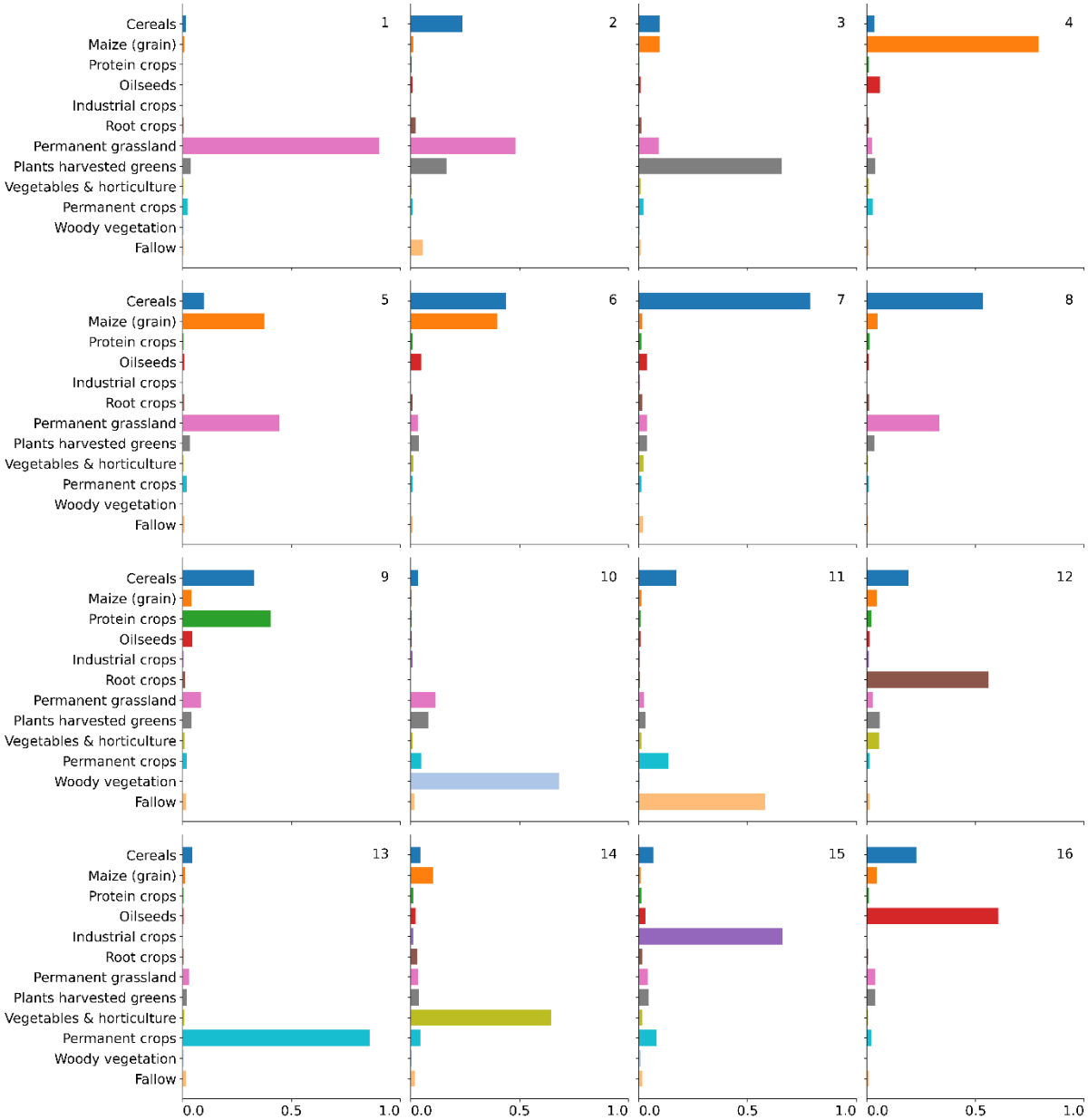


Figure 2: Relative crop composition portfolios of the 16 farm typologies identified using two-stage clustering of harmonised GSA data. Farm typologies broadly represent extensive production systems (Clusters 1 and 11), mixed grassland-cereal and

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

fodder-oriented systems (Clusters 2, 3, 5, and 8), cereal- and maize-dominated systems (Clusters 4, 6, and 7), specialised crop systems (Clusters 10, 13, 14, and 15), and diversified mixed farming systems (Clusters 9, 12, and 16).

3.2. Structural characteristics of farm typologies

The identified farm typologies also differ in structural farm characteristics, including mean farm size, number of fields, and cluster prevalence (Figure 3). Permanent grassland and cereal-dominated systems are among the most prevalent typologies and are generally characterised by medium to large farm sizes and moderate to high field numbers. This pattern is particularly evident in the permanent grassland clusters (Cluster 1 and 2) and the highly specialised cereal clusters (Cluster 6 and 7), while the more mixed cereal–grassland cluster (Cluster 8) shows relatively smaller farm sizes.

Other systems are characterised by smaller farm sizes and fewer fields. This is observed particularly in more specialised systems such as the maize-dominated systems (Cluster 4). The same pattern is evident in vegetable and horticultural farms (Cluster 14), as well as in mixed grassland-maize systems (Cluster 5), plants harvested green systems (Cluster 3), and industrial crop systems (Cluster 15), where farm sizes are generally small despite more moderate field numbers. Permanent crop systems (Cluster 13) represent a large number of farms and is likewise characterised by smaller farm sizes and moderate field numbers.

In contrast, clusters dominated by woody vegetation (Cluster 10), root crops (Cluster 12), fallow land (Cluster 11), protein crops (Cluster 9), and oilseeds (Cluster 16) generally have the largest farm sizes and moderate to high field numbers. Interestingly, farms dominated by fallow have relatively large farm sizes and a large number of individual fields.

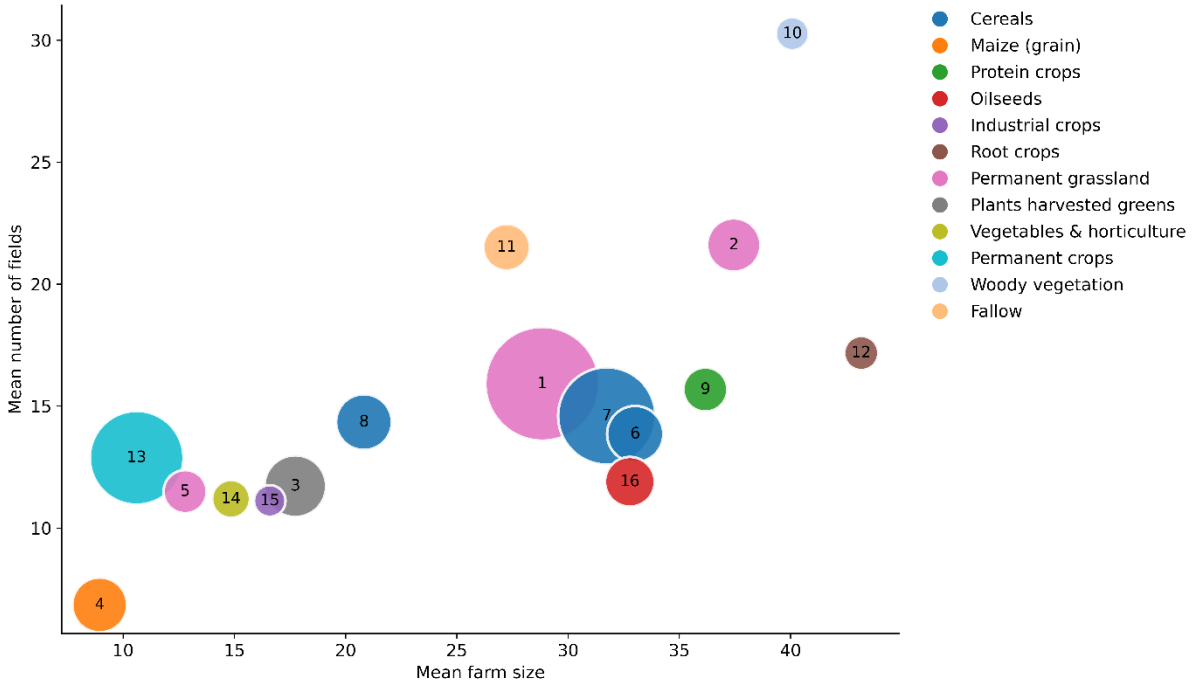


Figure 3: Structural characteristics of the identified farm typologies. Bubbles represent the dominant crop groups in each typology, while bubble size corresponds to the number of farms within each typology.

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

3.3. Spatial distribution of farm typologies

The spatial distribution of farm typologies shows clear regional differences across the study area (Figure 4). Permanent grassland typologies (Clusters 1, 2, and 5) are widespread across the study area, with particularly high concentrations in Ireland, northern and western Spain, Austria, and parts of Eastern Europe, especially Romania. In contrast, cereal-dominated typologies (Clusters 4, 6, 7, and 8) are broadly distributed across major arable farming regions in Central and Eastern Europe, particularly in Germany, Poland, and Romania, with additional concentrations in Spain and Denmark.

Mixed farming systems (Clusters 9, 12, and 16) are distributed across several regions, particularly in Romania, parts of Germany, Estonia, Spain, Italy, and the Netherlands. Fallow-dominated systems (Cluster 11) are strongly concentrated in Spain, while typologies dominated by plants harvested green (Cluster 3) are more widely distributed but especially prevalent in Denmark, northern Germany, and the Netherlands.

Industrial crop typologies (Cluster 15) are mainly concentrated in Italy and parts of Spain. Permanent crop (Cluster 13), woody vegetation (Cluster 10), and vegetable and horticultural typologies (Cluster 14) show stronger regional concentrations. Woody vegetation systems are primarily located in Portugal and Italy but also occur in parts of Denmark. Permanent crop systems are concentrated in Mediterranean regions such as Spain, Italy, Portugal, and Cyprus, while vegetable and horticultural systems are distributed across smaller but locally concentrated areas, particularly in Spain, Portugal, the Netherlands, and Belgium.

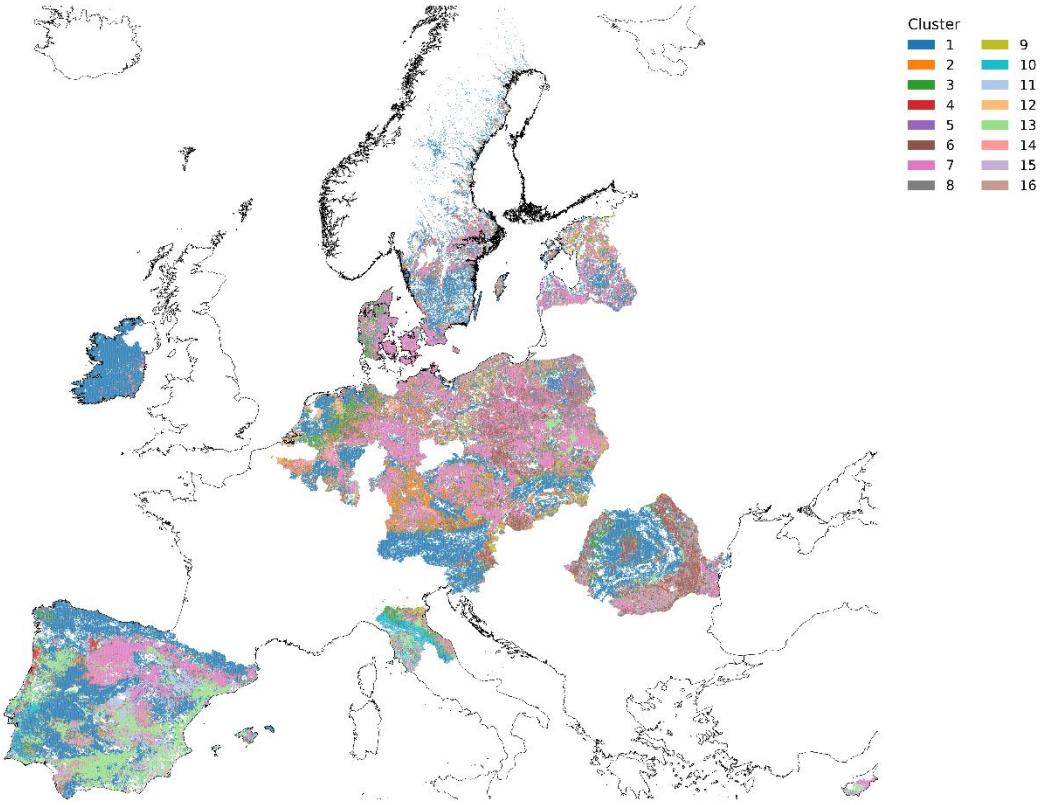


Figure 4: Spatial distribution of the 16 farm typologies derived from harmonised GSA data across 18 EU Member States.

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

4. Concluding remarks

The results demonstrate that harmonised GSA datasets can support the construction of meaningful cross-country farm typologies despite differences in national reporting systems and crop classification approaches. The identified typologies capture a broad range of production systems, including highly specialised crop systems, mixed farming systems, and variations in production intensity. In addition to differences in crop composition portfolios, the typologies also reflect variation in structural characteristics such as farm size and field counts, indicating that the clustering approach captures broader patterns of farm structure across the harmonised datasets.

The mapped typologies further revealed distinct regional distribution patterns, with several production systems strongly concentrated in specific agricultural regions, while others were more broadly distributed across the study area.

The study provides a transparent and replicable framework for constructing farm typologies from harmonised GSA datasets using harmonised crop composition portfolios within a two-stage clustering approach. At the same time, the study emphasises the importance of careful interpretation of classification-related uncertainties. The resulting maps and typologies should therefore be interpreted as harmonised analytical representations rather than fully standardised descriptions of agricultural systems across EU Member States.

While the present study covers 18 EU Member States, the resulting farm typologies are expected to become increasingly comprehensive and meaningful as more EU Member States provide GSA datasets containing detailed field-level agricultural land-use information and unique farm identifiers.

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

5. References

Beckmann, M., Didenko, G., Bullock, J. M., Cord, A. F., Paulus, A., Ziv, G., & Václavík, T. (2022). Archetypes of agri-environmental potential: A multi-scale typology for spatial stratification and upscaling in Europe. Environmental Research Letters, 17(11), 115008. https://doi.org/10.1088/1748-9326/ac9cf5

Erb, K.-H., Haberl, H., Jepsen, M. R., Kuemmerle, T., Lindner, M., Müller, D., Verburg, P. H., & Reenberg, A. (2013). A conceptual framework for analysing and measuring land-use intensity. Current Opinion in Environmental Sustainability, 5(5), 464–470. https://doi.org/10.1016/j.cosust.2013.07.010

Jepsen, M. R., Kuemmerle, T., Müller, D., Erb, K., Verburg, P. H., Haberl, H., Vesterager, J. P., Andrič, M., Antrop, M., Austrheim, G., Björn, I., Bondeau, A., Bürgi, M., Bryson, J., Caspar, G., Cassar, L. F., Conrad, E., Chromý, P., Daugirdas, V., … Reenberg, A. (2015). Transitions in European land-management regimes between 1800 and 2010. Land Use Policy, 49, 53–64. https://doi.org/10.1016/j.landusepol.2015.07.003

Jänicke, C., Petersen, K. A., Schmidts, P., Müller, D., & Jepsen, M. R. (2025). Field and farm-level data on agricultural land use for the European Union. Scientific Data, 12(1), 1050. https://doi.org/10.1038/s41597-025-05210-6

Jänicke, C., Petersen, K. A., Schmidts, P., Müller, D., & Jepsen, M. R. (2026). Harmonized IACS inventory (Version 1.3) [Dataset]. Zenodo. https://doi.org/10.5281/ZENODO.14230620

Jänicke, C., Wesemeyer, M., Chiarella, C., Lakes, T., Levers, C., Meyfroidt, P., Müller, D., Pratzer, M., & Rufin, P. (2024). Can we estimate farm size from field size? An empirical investigation of the field size to farm size relationship. Agricultural Systems, 220, 104088. https://doi.org/10.1016/j.agsy.2024.104088

Kohonen, T. (2001). Self-Organizing Maps (Bd. 30). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-56927-2

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., Plutzar, C., Verkerk, P. J., Van Der Zanden, E. H., & Reenberg, A. (2013). Challenges and opportunities in mapping land use intensity globally. Current Opinion in Environmental Sustainability, 5(5), 484–493. https://doi.org/10.1016/j.cosust.2013.06.002

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. Oecologia, 129(2), 271–280. https://doi.org/10.1007/s004420100716

Leonhardt, H., Wesemeyer, M., Eder, A., Hüttel, S., Lakes, T., Schaak, H., Seifert, S., & Wolff, S. (2024). Use cases and scientific potential of land use data from the EU's Integrated Administration and Control System: A systematic mapping review. Ecological Indicators, 167, 112709. https://doi.org/10.1016/j.ecolind.2024.112709

Satopaa, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011). Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior. 2011 31st International Conference on Distributed Computing Systems Workshops, 166–171. https://doi.org/10.1109/ICDCSW.2011.20

Schneider, M., Schelte, T., Schmitz, F., & Körner, M. (2023). EuroCrops: The Largest Harmonized Open Crop Dataset Across the European Union. Scientific Data, 10(1), 612. https://doi.org/10.1038/s41597-023-02517-0

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. IEEE Transactions on Neural Networks, 11(3), 586–600. https://doi.org/10.1109/72.846731

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, 58(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Funded by the European Union

Funded by the European Union (101081307). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EC-CINEA. Neither the European Union nor the granting authority can be held responsible for them.