AGRICULTURAL AND ENVIRONMENTAL MODELLING
Publishing data and modelling research objects

Agricultural and Environmental Modelling 8: e205420 (2026)
DOI: 10.3897/aem.8.205420

Data Paper (Biosciences)

High-resolution agricultural landscape models for selected agricultural regions across six European countries for use in the ALMaSS simulation framework

Elżbieta Ziółkowska[1,2,3], Geoff Groom[3], Jean-Sébastien Beaulne[4], Anto Raja Dominic[5], Hella Kehlenbeck[5], Luna Kondrup Marcussen[3], Łukasz Mikołajczyk[1], Anne Muola[6], Sari Peltonen[7], Benoit Sarrazin[8], Jürgen Schwarz[5], Susanne Stein[9], Bettina Wenzel[5], Christopher John Topping[3]

1 *Institute of Environmental Sciences, Department of Biology, Jagiellonian University, Kraków, Poland*
2 *Institute of Geography and Spatial Management, Department of Geography and Geology, Jagiellonian University, Kraków, Poland*
3 *Social-Ecological Systems Simulation Centre, Department of Agroecology – Agricultural biodiversity, Aarhus University, Aarhus, Denmark*
4 *Institut Camille Jordan, Lyon 1 Université Claude Bernard, Villeurbanne, France*
5 *Institute for Strategies and Technology Assessment, Julius Kühn-Institut, Kleinmachnow, Germany*
6 *Division of Biotechnology and Plant Health, Norwegian Institute of Bioeconomy Research, Tromsø, Norway*
7 *ProAgria, Vantaa, Finland*
8 *Unité Agroécologie et Environnement, ISARA, Lyon, France*
9 *Institute of Chemistry and Biochemistry, Department of Biology, Chemistry, and Pharmacy, Freie Universität Berlin, Berlin, Germany*

Corresponding author: Elżbieta Ziółkowska (e.ziolkowska@uj.edu.pl)

Abstract

This data paper presents a harmonized set of spatial datasets describing 60 agricultural landscapes (10 × 10 km each) generated for selected agricultural regions across six European countries—Denmark, Finland, France, Germany, the Netherlands, and Poland—within the Horizon 2020 EcoStack project. These landscape datasets were developed for use within the Animal, Landscape and Man Simulation System (ALMaSS) and applied within an integrated modeling framework designed to identify management strategies that effectively enhance multiple beneficial organisms, i.e., ecosystem service providers, within European agroecosystems. All landscape models were generated following the methodology detailed in Ziółkowska et al. (2026). Because input data availability and structure varied among countries, the full set of data sources and the processing steps required to account for national and regional specificities are documented. All inputs were harmonized to a common structure, spatial resolution, and Type of Landscape Element (TOLE) legend to support comparability across landscapes. The resulting datasets capture major gradients in farmland structure, composition, and management intensity across Europe. They can support not only agent-based simulations and scenario analyses but also ecological modeling and biodiversity assessments more broadly. All landscape models, associated files, metadata, and landscape-window selection criteria are provided in an open, citable format to facilitate reuse and further development.

Key words: Agroecology, data harmonization, dynamic landscape models, land use and land cover, landscape structure, spatio-temporal heterogeneity

OPEN ACCESS

Academic editor: Gowhar Meraj
Received: 26 June 2026
Accepted: 2 October 2026
Published: 4 November 2026

Citation: Ziółkowska E, Groom G, Beaulne J-S, Dominic AR, Kehlenbeck H, Marcussen LK, Mikołajczyk Ł, Muola A, Peltonen S, Sarrazin B, Schwarz J, Stein S, Wenzel B, Topping CJ (2026) High-resolution agricultural landscape models for selected agricultural regions across six European countries for use in the ALMaSS simulation framework. Agricultural and Environmental Modelling 8: e205420. https://doi.org/10.3897/aem.8.205420

Copyright: © Elżbieta Ziółkowska et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

Agricultural expansion and intensification have reduced and fragmented semi-natural habitats across Europe, affecting both animal and plant communities and their interactions. Many non-crop species are increasingly confined to habitat remnants such as field margins, hedgerows, and small semi-natural patches embedded in intensively farmed landscapes (e.g., Tscharntke et al. 2005; Priyadarshana et al. 2024). These elements act as essential biodiversity reservoirs and support key ecosystem services, including pollination and natural pest control. However, their effectiveness varies across landscape contexts, management histories, and local ecological conditions (Fahrig et al. 2010; Martínez-Núñez et al. 2022). Semi-natural habitats benefit both pollinators (through floral resources) and natural enemies (through alternative prey, overwintering sites, and shelter), although the relative importance of specific habitat types differs between these functional groups. Agri-environment measures that create or maintain such habitats can, in principle, support multiple ecosystem service providers (ESPs) simultaneously. More broadly, combining beneficial services of functional biodiversity from all levels—from soil microbes and beneficial arthropods to vertebrates—in a synergistic manner to maximize ecosystem service delivery is referred to as ecostacking (Hokkanen 2017). Yet species and functional groups respond differently to individual interventions, and combinations of measures often have non-additive outcomes on the abundance and diversity of ESPs. Therefore, it remains difficult to identify the management strategies that most efficiently enhance multiple taxa of ESPs simultaneously. Thus, well-defined management goals are necessary for designing and evaluating ecostacking strategies.

Integrative modeling approaches capable of predicting how alternative management practices affect ESPs across diverse European agricultural systems have been developed in recent years (Chetcuti et al. 2024). This has been achieved by coupling detailed species-specific behavioral and ecological models with a dynamic landscape simulation within the Animal, Landscape and Man Simulation System (ALMaSS; Topping and Hansen 2003; Topping 2022). The landscape simulation in ALMaSS allows for capturing both the spatial structure of agricultural environments (defined through a high-resolution land-use/land-cover map including farm locations and characteristics) and their temporal dynamics, reflecting crop- and field-level management operations and their effects on vegetation development of both crops and semi-natural elements. These operations are closely linked to models of vegetation biomass growth and floral resource availability, which are driven by weather inputs and vary by geographic location, farming system, and habitat type. Operating on daily time steps, ALMaSS allows individuals of the modeled focal species to interact with a landscape that changes continuously through vegetation growth, farming activities, and weather-dependent fluctuations in both abiotic conditions and biotic resources. Variation in the composition and spatial arrangement of cropping systems emerges from farm structure, farm-type-specific crop rotations, and the underlying crop and vegetation models (Topping and Duan 2024; Ziółkowska et al. 2026).

For ALMaSS simulations to be ecologically meaningful and comparable across countries, the underlying landscape models must meet demanding

standards: they must capture the spatial heterogeneity of agricultural land at field resolution, accurately represent semi-natural habitat types and their connectivity, and encode realistic farm-level management, including crop rotations and agrochemical use. Producing such models requires integrating diverse national datasets (topographic databases, land parcel identification systems, agricultural registers, soil maps, and weather records) into a consistent, harmonized structure. A unified methodology for generating 10 × 10 km ALMaSS landscape models from these sources was developed and is described in detail in Ziółkowska et al. (2026). It was implemented in selected agricultural regions across six European countries—Denmark, Finland, France, Germany, the Netherlands, and Poland—chosen to represent the major biogeographical regions of Europe under the EU Habitats Directive (92/43/EEC): the Atlantic (Denmark, Netherlands), Continental (Germany, Poland), Southern Continental (France), and Boreal (Finland) regions. Ten landscape windows per country were selected, generating 60 landscape models in total, spanning a broad gradient of agricultural structures, farming systems, and semi-natural habitat configurations across the continent.

Here, the complete set of ALMaSS-ready input files for these 60 landscape windows is presented. The models were originally developed to support scenario analyses of agri-environment measures targeting multiple ESPs within the Horizon 2020 EcoStack project and have been applied to simulate the responses of generalist predators (*Bembidion lampros*, *Erigone atra*), wild pollinators (*Osmia bicornis*), a ladybird–aphid predator–prey system (*Coccinella septempunctata* and four aphid species), Eurasian skylark (*Alauda arvensis*), European hare (*Lepus europaeus*), and field vole (*Microtus agrestis*) to landscape-scale management interventions. However, the dataset is not limited to these applications. Because each landscape model captures a high-resolution, spatially and temporally explicit representation of agricultural land use, farm structure, and semi-natural habitat, it can support a wide range of ecological modeling, biodiversity assessment, and landscape analysis tasks beyond agent-based simulation.

For each country or region, the data sources used, criteria for selecting landscape windows, the country- or region-specific processing steps required to harmonize heterogeneous national datasets, and the key structural characteristics of the resulting landscapes are documented. The 60 landscape models described here form part of a broader collection of ALMaSS landscape models produced through multiple European and national research initiatives.

Methods for data generation

The spatial heterogeneity in the ALMaSS landscape model is represented through a detailed land-use/land-cover map (hereafter the ALMaSS landscape map). To optimize memory use, the system combines raster and patch-based representations (Topping and Duan 2024): each raster cell stores a reference to a corresponding homogeneous polygon (patch) of a given land-use or land-cover type. These polygons hold the full set of descriptive attributes: geometric properties, soil type, elevation, farm type, and other biotic or abiotic characteristics. This flyweight-based structure (Gamma et al. 1994) minimizes duplication of identical polygon information and enables efficient storage even

for large landscapes. The raster map can represent areas of up to 50 × 50 km, although landscape windows of 10 × 10 km at 1 m^2 resolution are typically used in ALMaSS simulations.

The temporal component allows polygon properties in ALMaSS simulations to vary daily, reflecting vegetation dynamics (height, green and total biomass) of both crop and semi-natural vegetation, floral resources (modeled as daily pollen, nectar, and sugar availability, parameterized from plant species assemblages characteristic of each habitat type), and responses to weather and farm management. Crop and vegetation changes follow crop-specific management plans and multi-year crop rotations, enabling the landscape to evolve realistically throughout the farming year.

The ALMaSS landscape model is defined through a set of input files that support landscape simulation via the Landscape and FarmManager C++ classes. The Landscape class creates the internal representation of the ALMaSS landscape map, along with the supporting classes needed to efficiently manage spatial data and relationships. The FarmManager class simulates cropping systems and crop husbandry operations. Details of these processes are provided in Topping and Duan (2024).

Below, an overview of the types and structure of input files used in ALMaSS landscape simulations is first provided. The input data used for landscape-model generation in the six selected countries are then summarized, and the country-specific data-processing procedures, which differ in several aspects from the general workflow outlined in Ziółkowska et al. (2026), are described.

Input files for landscape simulation in ALMaSS

Input files for ALMaSS landscape simulation are divided into landscape-specific and country- or region-specific categories (Table 1). Landscape-specific files describe the spatial structure and local weather for a given simulation window. Spatial structure is encoded in three components: the binary raster landscape map (*.lsb*), a polygon reference file (*PolyRef*), and a farm reference file (*FarmRef*). The *.lsb* file defines the landscape's spatial extent and resolution. It assigns each raster cell a polygon reference number (PolyRefNum) corresponding to the polygon that occupies the largest proportion of that cell. The *PolyRef* file provides, for every polygon reference number, the associated polygon attributes, e.g., type of landscape element (TOLE), soil type, elevation, slope, aspect, and links agriculturally managed polygons to their respective farms via the *FarmRef* reference number. The *FarmRef* file, in turn, lists all farms within the landscape and specifies their management type (e.g., conventional arable, organic cereal), enabling connections between farms, polygons, and the landscape map. This linkage is essential because, within ALMaSS landscape simulation, agricultural areas are managed at the farm level: a Farm object (instantiated from the Farm class) represents each farm unit, allocating crops and executing management activities in each field associated with that farm (Topping and Duan 2024).

Weather inputs are also landscape-specific. Each landscape is associated with a *.pre* text file containing hourly historical records of temperature, precipitation, wind speed and direction, soil temperature, snow depth, humidity, and

Table 1. Summary of input files for landscape simulation in ALMaSS. * Crop management plans are not provided as external files but are implemented in the ALMaSS code. Supporting Excel files facilitate the translation of management information into model code. ** All vegetation growth parameters for all vegetation types across all countries/regions are included in one *.pre* text file. *** All floral resource curves for all habitats across all countries/regions are included in one text file.

Input file	Category	Content/description	Purpose in ALMaSS
{landscape_name}.lsb	Landscape-specific	Binary raster map of the landscape; each cell references the polygon to which it belongs (via *PolyRefNum*).	Represents spatial structure of the landscape and links raster cells to polygon attributes.
{landscape_name}_PolyRef.txt	Landscape-specific	Contains attributes of polygons (soil type, elevation, slope, and aspect); links polygons to the landscape map via *PolyRefNum*; links polygons to farms via *FarmRef* reference number.	Provides polygon-level details for the landscape map and enables connection to farm units.
{landscape_name}_FarmRef.txt	Landscape-specific	Lists farm units in the landscape and their types (e.g., conventional arable and organic cereal); links to *{landscape_name}_PolyRef.txt* via *FarmRef* reference number.	Links polygons to farm units and defines farm types.
{landscape_name}.pre (weather)	Landscape-specific	Hourly historical weather data (temperature, precipitation, wind, soil temperature, snow depth, humidity, and radiation); typically 10-year records	Drives temporal dynamics in ALMaSS, including crop growth, floral resources, and other weather-dependent processes.
{farm_type}.rot	Country/ region-specific	Farm-type-specific crop rotations (sequence of crops over multi-year rotations).	Allows the assignment of crops to fields during landscape simulation in accordance with farm type and ensures realistic crop proportions and temporal patterns.
Crop management plans*	Country/ region-specific	Sequence, timing, and probability of farming operations (soil cultivation, fertilization, and pesticide application); embedded in ALMaSS code.	Controls crop management and husbandry operations in simulations.
.pre (vegetation growth)**	Country/ region-specific	Daily growth models for all vegetated elements, including crops; links to vegetation types (TOV—types of vegetation) via reference numbers.	Simulates vegetation dynamics (height, biomass, and other characteristics).
.txt (floral resources)**	Country/ region-specific	Daily floral resource models for all habitats associated with vegetated elements, including crops. Links to *{landscape_name}_PolyRef.txt* via PollenNectarCurve reference number.	Simulates floral resource dynamics (amounts of pollen, nectar, and sugar produced daily).

radiation. Typically, 10 years of data are used and looped for longer simulations (Topping and Duan 2024).

Country- or region-specific inputs include farm-type crop rotations stored as .rot files. These files define the crop sequence for each farm type. Permanent crops maintain fixed vegetation types, whereas fields in rotation are assigned crops according to the farm type, first randomly at simulation start and then following the defined rotation sequence. Rotations typically include 100 entries to reflect observed crop proportions. All crops must have associated crop management plans that define sequences, timing windows, and the probabilities of operations such as soil cultivation, fertilization, and pesticide application. These plans are embedded in the ALMaSS code and parameterized from regional or national sources, including farmer interviews and official crop production guidelines; supporting Excel files facilitate translation into the ALMaSS code.

Vegetation growth models, required for daily simulation of vegetation dynamics, are defined in a single *.pre* file for all vegetated landscape elements, including crops. This file contains growth parameters for all vegetation types across all countries/regions, and internal ALMaSS mechanisms link vegetation types to the appropriate model via reference numbers. Similarly, floral resource models that allow daily simulation of pollen, nectar, and sugar production are stored in a single text file. Each vegetated TOLE is assigned its own dedicat-

ed floral resource curve characterizing the seasonal profile of daily resource availability for that habitat type. These curves are constructed by aggregating species-level models—integrating resource production per floral unit, flowering abundance, and flowering phenology—weighted by each species' proportional cover within the characteristic plant community of that TOLE. The resulting habitat-level curves are linked to landscape polygons via a PollenNectarCurve reference number stored in the *PolyRef* file. The methodology underlying these floral resource models was developed within the Horizon 2020 B-GOOD and Horizon Europe Better-B projects. Although floral resource curves are the only habitat-quality metric stored at the polygon level, the landscape models support simulation of a wide range of species groups beyond pollinators, as habitat suitability for other taxa (including predatory arthropods, small mammals, and birds) is determined through species-specific functional classifications of TOLEs within the ALMaSS species models.

Together, these files constitute the core inputs that govern how farms, crops, vegetation, weather, and management processes interact within ALMaSS to produce spatially and temporally dynamic landscape simulations.

For landscape simulations to function correctly, all input files must be consistently aligned with the internal data structures defined in the ALMaSS code (Topping and Duan 2024). Crop types used in *.rot* rotation files must first be implemented as ALMaSS Types of Vegetation (TOVs) and associated with the appropriate crop management plans and vegetation growth models. Likewise, all land-cover and land-use categories in the raster landscape map must be correctly translated into ALMaSS Types of Landscape Elements (TOLEs) and be referenced accordingly in the *PolyRef* file. Introducing new TOLEs outside the existing set requires explicitly implementing them in the code. To enable meaningful interactions between landscapes and species models, TOLEs and TOVs must also be classified by their functional roles (e.g., breeding, foraging, movement habitat, or non-habitat). These linkages are essential for ensuring that landscape structure, management processes, and species ecology operate coherently. Systematic testing is therefore required whenever new national or regional landscape models are added to ALMaSS to verify that all structural, management, and ecological connections are correctly established.

Overview of input data used to generate ALMaSS landscape models in selected countries

The landscape models presented in this paper were generated using the general workflow described in detail in Ziółkowska et al. (2026), implemented via semi-automatic Python scripts adapted for each country or region. New landscape models were developed for Poland, Finland, and France (restricted to the Auvergne–Rhône–Alpes region due to data availability) within the EcoStack project. For Denmark, earlier ALMaSS landscape models were updated within the same project. The models for the Netherlands were developed within a project funded by the PBL Netherlands Environmental Assessment Agency (Ziółkowska et al. 2022), and the models for Germany (Brandenburg and Lower Saxony federal states) were developed under the ELONTA project funded by the German Environment Agency (UBA) (Ziółkowska et al. 2024).

Land cover/land use and topography

ALMaSS landscape maps are derived from spatial information describing land use and land cover, such as water networks, transport networks (including roads and railways), buildings, and other infrastructure elements. All important natural or semi-natural landscape features (e.g., rivers or streams, lakes, and bare sand or rock) and forms of semi-natural vegetation (e.g., hedgerows, open vegetation types, forest, and isolated trees) should be adequately represented (Ziółkowska et al. 2026). The primary input for generating these maps consists of national topographic object databases, which provide the necessary geometric and thematic detail corresponding to map scales between 1:5,000 and 1:25,000; these datasets for the six countries included in this study are summarized in Table 2. Depending on the country, these core topographic datasets can be supplemented by additional sources that supply specific information not represented, or insufficiently detailed, in the national databases. Such auxiliary datasets may include additional layers besides agricultural parcels from the Land Parcel Identification System (LPIS) (e.g., to map small woody features within agricultural land in Germany or managed grasslands in Poland) or satellite-derived products (e.g., National Forest Inventory and CORINE Land Cover in Finland) to separate certain land-use/land-cover categories and fill in missing information about open land (Table 2).

Because the land-cover/land-use and topography input data differ across countries in terms of thematic resolution, definitions of mapped features, and data structure, producing harmonized ALMaSS landscape maps requires country-specific processing. This includes extraction of relevant thematic classes, transformation of heterogeneous attribute schemas, and integration of supplementary datasets to achieve a comparable level of detail across modeling regions despite differences in national mapping standards.

Agriculture

Agricultural areas and their fine-scale spatial heterogeneity in land cover and management are particularly important for the ALMaSS landscape model and therefore require detailed mapping and characterization. This necessitates not only the delineation of individual fields but also their allocation to the appropriate farm units, with each unit further characterized by its management type and associated crop rotations. Such high-resolution spatial and temporal agricultural information is available through the Integrated Administration and Control System (IACS), the EU-wide control framework underpinning the Common Agricultural Policy (CAP). One of the core components of IACS is the Land Parcel Identification System (LPIS), which stores geospatial information on reference parcels—geographically delineated units with unique identification codes. Additional IACS components record annual information on crops grown on each reference parcel (direct payment registers), maintain agricultural holding registers that allow parcels to be grouped into farm units, and store data on livestock kept for agricultural purposes. Because IACS is administered at national or regional levels, update frequency and detail typically follow an annual cycle but may vary among Member States. By integrating these compo-

Table 2. Datasets used for generating ALMaSS landscape maps in different countries. The table lists the core topographic databases for each country and additional thematic datasets required for achieving comparable landscape representation of non-agricultural areas across countries. Data used for mapping the structure of agricultural areas are presented in Table 3.

Country	Dataset	Description	Source/data provider	Data currency (year)	Format	Scale/resolution	Additional information
Denmark	GeoDanmark (GeoDK)	Topographic database	Agency for Data Supply and Efficiency (SDFE)	2020	Vector	1:10,000	Base source data
	KORT10	Topographic database	National Survey and Cadastre of Denmark	2013	Vector	1:10,000	To map copses, pylons, and single trees
	Fund og Fortidsminder	Geodatabase of archeological sites	Danish Castles and Culture Agency	2020	Vector	n/a	To map burial mounds (refuges for wildlife)
	Digital Terrain Model (DTM) and Digital Surface Model (DSM)	National airborne laser scanning data	Agency for Data Supply and Efficiency (SDFE)	2014/2015	Raster	0.4 m	To map slopes alongside roads, rail lines, and hedges
Finland	Maastotietokanta	Topographic database	National Land Survey of Finland (NLS)	2019	Vector	1:5,000–1:10,000	Base source data
	National Forest Inventory (NFI)	Forest canopy cover	Natural Resources Institute Finland	2011	Raster	20 m	To map forest patches
	Corine Land Cover (CLC)	Land cover database of Europe	Copernicus LAND	2018	Vector	MMU = 25 ha; MMW = 100 m	To map mining areas
Germany	Basis-DLM	Topographic database	Authorative Topographic-Cartographic Information System (ATKIS)	2019	Vector	1:25,000	Base source data
	BKG	Buildings	Federal Agency for Cartography and Geodesy	2016	Vector	n/a	To map buildings
	LPIS	Land parcel identification system.	Brandenburg (LELF) and Lower Saxony (LWK) Federal States	2019	Vector	n/a	To map hedgerows and tree lines
France	BD TOPO	Topographic database	Institut national de l'information géographique et forestière (IGN-F)	2019	Vector	1:5,000–1:10,000	Base source data
	OSCOM	Topographic database	Direction Régionale de l'Environnement de l'Aménagement et du Logement d'Auvergne-Rhône-Alpes (DREAL Auvergne-Rhône-Alpes)	2019	Vector	1:10,000	To map additional vegetation features (e.g., poplar forests, woody moors, meadows, or forage areas)
Poland	BDOT	Topographic database	Head Office of Geodesy and Cartography (GUGiK)	2019	Vector	1:10,000	Base source data
	LPIS	Land Parcel Identification System	Polish Agency of Restructuring and Modernization of Agriculture	2018	Vector	n/a	To map managed permanent grasslands
Netherlands	TOP10NL	Topographic database	Land Registry and Mapping Agency	2017	Vector	1:5,000–1:25,000	Base source data

nents, agricultural land in the ALMaSS landscape map can be subdivided into individual fields and farm units (*PolyRef* file), and farm units can be classified into management types (*FarmRef* file) based on the combination of crops cultivated and livestock raised.

Although the LPIS reference parcel layer is a mandatory element of EU subsidy schemes—requiring farmers to delineate their parcels for payment claims—the level of spatial and thematic detail provided differs among countries and regions. These differences stem partly from the range of reference parcel types permitted under EU legislation. Article 6.1 of Commission Regulation No. 796/2004 allows four reference parcel models: cadastral parcels, agricultural parcels, farmers' blocks, and physical blocks. LPIS implementations based on

cadastral parcels present particular challenges because cadastral units reflect property ownership rather than land use, whereas the other LPIS models are designed explicitly around land-use or land-cover criteria. Among the six countries considered, five used agricultural-parcel-based LPIS systems at the time of data acquisition, while Poland relied on a cadastral-based system (an agricultural-parcel-based LPIS system was implemented in Poland in 2020). This required additional processing because cadastral parcels often contain multiple land-use or land-cover classes (e.g., farmhouses, gardens, forest patches, and streams) alongside agricultural fields. Moreover, the agricultural register reporting crop types was also referenced to cadastral parcels; thus, when a parcel contained more than one agricultural field, the register provided only crop areas at the parcel level, without field-specific boundaries. Consequently, an extra reconstruction step was necessary for Poland to delineate individual fields before integrating them into the ALMaSS landscape maps.

In addition, LPIS does not necessarily cover all agriculturally managed areas, as not all farmers submit subsidy claims. The ALMaSS landscape generation workflow (Ziółkowska et al. 2026) incorporates procedures to account for such omissions when defining appropriate data processing steps. However, the magnitude of this issue may vary considerably among countries and regions. Areas dominated by small (<1 ha), heterogeneous fields (common in many agricultural regions in Poland) are especially challenging, as omissions in LPIS coverage may disproportionately affect the accurate delineation and classification of agricultural land in such landscapes.

Another challenge arises with regard to access to IACS data. While Regulation (EU) 2023/138 requires high-value geospatial datasets, including LPIS/GeoSpatial Aid Application (GSAA), to be made available free of charge, in machine-readable form, and under open reuse licenses (e.g., CC-BY 4.0 or equivalent), the actual level of openness depends on how each Member State implements the requirements and reconciles them with data-privacy legislation such as the General Data Protection Regulation (GDPR). As a result, full, raw IACS data are often restricted; publicly available versions in many countries provide only basic crop information per parcel. To obtain anonymized reference-parcel owner/applicant identifiers or holding-unit identifiers (required to group parcels into farm units), formal requests must be submitted to the responsible national or regional data providers. The situation is further complicated in countries where IACS is managed at a regional or state level, for example, in Germany or France, because each region or state must be contacted individually. In this study, these access restrictions resulted in different spatial extents and types of data obtained from IACS among the six countries (see Table 3). Specifically, in the case of countries with IACS data managed at the regional level, data were obtained for two German federal states (Brandenburg and Lower Saxony) and for the French region Auvergne–Rhône–Alpes. In addition, in the French case, livestock data necessary for farm type classification could not be obtained, but herd-location data were acquired, which served as a substitute in the farm classification process.

In ALMaSS, each crop modeled, as either conventional or organic, is associated with a crop management that captures its characteristic management regime. These regimes originate from detailed crop management plans describing the sequence, timing, and likelihood of key farming activities such as soil

Table 3. Datasets used for generating ALMaSS landscape maps in different countries. The table lists data used for mapping the structure of agricultural areas, generating farm classifications, and generating crop rotations.

Country	Dataset	Description	Source/data provider	Data currency (year)	Format	Reference unit
Denmark	General Farm Register (Det Generelle Landbrugsregister – GLR)	Field boundaries, area, and type of crops cultivated, together with the ID of agricultural holdings	Ministry of Food, Agriculture and Fisheres of Denmark (FVM)	2019	Vector	Field
	Central Livestock Register (Det Centrale HusdyrbrugsRegister; CHR)	Number of animals	Danish Center for Environment (DCE)	2019	Tabular	Agricultural holding
	Crop management plans	Descriptions of occurrence, timing, and probability of farming activities	Farmer advisors	2013–2020	Tabular	Crop
Finland	LPIS	Field boundaries, area, and type of crops cultivated, together with the ID of agricultural holdings	Finnish Food Authority	2018	Vector	Field
	Animal identification system	Number of animals	Finnish Food Authority	2018	Tabular	Agricultural holding
	Crop management plans	Descriptions of occurrence, timing, and probability of farming activities	Proagria	2019–2020	Tabular	Crop
France	RPG (Registre Parcellaire Graphique)	Field boundaries, area, and type of crops cultivated, together with the ID of agricultural holdings	Agence de services et de paiement (ASP)	2019	Vector	Field
	Lieux anonymes de détention d'animaux	Location of animals' herds	Ministry of Agriculture and Food (Direction générale de l'Alimentation; DGAL)	2019	Vector	Herd
	Crop management plans	Descriptions of occurrence, timing, and probability of farming activities	ISARA	2019–2020	Tabular	Crop
Germany	LPIS	Field boundaries, area, and type of crops cultivated, together with the ID of agricultural holdings	Brandenburg (LELF) and Lower Saxony (LWK) Federal States	2019	Vector	Field
	Animal data	Number of animals	Brandenburg (LELF) and Lower Saxony (LWK) Federal States	2019	Tabular	Agricultural holding
	Crop management plans	Descriptions of occurrence, timing, and probability of farming activities	Julius Kühn Institute	2022	Tabular	Crop
Poland	National land and building cadaster (LBC)	Reference parcels for LPIS	Polish Agency of Restructuring and Modernization of Agriculture	2018	Vector	Cadastral parcel
	Agricultural register	Area and type of crops cultivated, together with the ID of agricultural holdings	Polish Agency of Restructuring and Modernization of Agriculture	2018	Tabular	Cadastral parcel
	Animal Identification and Registration System (AIRS)	Number of animals	Polish Agency of Restructuring and Modernization of Agriculture	2018	Tabular	Agricultural holding
	Crop management plans	Descriptions of occurrence, timing, and probability of farming activities.	Agricultural Advisory Centre, regional offices	2020	Tabular	Crop
Netherlands	Basisregistratie Percelen	Field boundaries, area, and type of crops cultivated, together with the ID of agricultural holdings and farm classification	Netherlands Enterprise Agency (Rijksdienst voor Ondernemend Nederland)	2015	Vector	Field
	Crop management plans	Descriptions of occurrence, timing, and probability of farming activities	Delphy	2018	Tabular	Crop

cultivation, fertilizer use, and pesticide applications. The plans are developed using information gathered through interviews with farmers and farm advisors, as well as national guidelines on integrated crop production. Incorporating time windows and operational probabilities allows ALMaSS to simulate the spatial and temporal organization of farming practices and the resulting changes in vegetation structure at the landscape scale.

Preparing crop management plans at this level of detail is time-consuming, so the number of crop types modeled in each country or region must be bal-

anced to adequately represent crop heterogeneity while remaining feasible in terms of data collection. Typically, only crops covering more than 1% of the agricultural area within each farm type are selected, provided that obtaining reliable management information is possible. Representative crops from major crop groups (such as winter and spring cereals, fodder crops, and vegetables) are also included to ensure that realistic crop rotations can be generated within ALMaSS. Consequently, across the six countries included in the EcoStack project, both the number of crop management plans and the data sources from which they were compiled varied (Table 3).

Additional information layers: soil

The ALMaSS landscape simulator uses the dominant soil type at the field level to parameterize farming operations that rely on soil conditions; soil maps are therefore required to assign a soil class to each polygon in the landscape map, recorded in the *PolyRef* file. ALMaSS uses a 15-class soil classification derived from the FAO soil-texture system and linked to agricultural use and soil-management categories (Topping and Duan 2024). Consequently, national or regional soil classes must be translated into the ALMaSS system, ensuring consistency across countries. Soil information may also be incorporated directly into crop-management plans to differentiate operations by soil type and can be further linked to polygon-level soil attributes where relevant. For the five countries included in this study, soil data were obtained from national soil maps wherever possible (Table 4). For Poland, due to access costs, data from the European Soil Database (https://esdac.jrc.ec.europa.eu/resource-type/datasets) provided by the Joint Research Centre were used.

Additional information layers: weather

Historical weather data are required to drive both vegetation growth and floral resource models. They also influence (and sometimes constrain) farming activities and drive the development and behavior of modeled animals within the ALMaSS system. For all 60 simulated landscapes, 10-year hourly weather inputs—including mean temperature, precipitation, wind speed, wind direction,

Table 4. Soil data sources used for the ALMaSS landscape model for each country.

Country	Soil data provider	Format	Scale/ resolution	Map name/remarks
Denmark	Aarhus University Department of Agroecology 2013	Raster	1:50,000/30.4 m	National high-resolution 3-D mapping of soil texture
Germany	German Federal Institute for Geosciences and Natural Resources (BGR)	Vector	1:200,000	National soil map: BUEK200
Finland	Geological Survey of Finland	Vector	1:200,000	National soil map
France	French Geological Survey (BRGM)	Vector	1:250,000	National soil map
Netherlands	National Soil Information System (BIS) managed by Wageningen Environmental Research on behalf of the Ministry of Agriculture, Fisheries, Food Security and Nature	Vector	1:50,000	High-resolution national soil map: Bodemkaart (SGM)
Poland	European Soil Database (JRC)	Vector	1:1,000,000	Used because national data in scale 1:100,000 (provided by the Institute of Soil Science and Plant Cultivation) is not free of charge at the time of acquisition (~0.5 euro/km^2)

soil temperature, snow depth, humidity, and radiation—were generated by extracting the corresponding parameter values from the ERA5 gridded dataset at the coordinates of the upper-left corner of each landscape window. ERA5 global hourly climate and weather data are provided by the Copernicus Programme (Hersbach et al. 2020). Details of how weather data are implemented in ALMaSS are described in Topping and Duan (2024).

Input data limitations

Source datasets varied in spatial resolution and scale as well as acquisition date (Tables 2–4). The core topographic and land-cover datasets from which the ALMaSS landscape maps were derived were of broadly comparable scale across the six countries (ranging from 1:5,000 to 1:25,000; Table 2), supporting the comparability of landscapes derived from these sources. Soil data are a notable exception, ranging from 1:50,000 in Denmark and the Netherlands to 1:1,000,000 in Poland (Table 4); however, soil type in ALMaSS is used only to parameterize crop management operations and is not itself reported or compared as a landscape structure metric, so this disparity in soil-data resolution does not affect the comparability of the landscape models presented. Woody vegetation and acquisition dates also warrant particular caution for Finland: whereas woody elements in the other five countries were delineated from vector-based topographic databases at comparable positional accuracy to the rest of the landscape map, Finnish forest cover was derived from National Forest Inventory canopy-cover rasters at 20 m resolution (2011), a coarser, threshold-based classification predating most other input layers (compiled approximately 2016–2020; Table 2). Finnish landscapes should therefore be compared to those of other countries with some caution, reflecting differences in the spatial accuracy and acquisition date of the underlying woody-vegetation data rather than a difference in mapping approach. A further temporal difference concerns the Netherlands, where the field-boundary and crop-type dataset (Basisregistratie Percelen) dates from 2015, three to four years earlier than the equivalent agricultural registers used for the other five countries (2018–2019; Table 3); more recent data were not available to the authors at the time of acquisition. However, comparison against coarser-scale national agricultural block statistics indicates that the spatial structure of field parcels and the relative shares of major crop types in the Netherlands have remained broadly stable since 2015, suggesting that this dataset continues to provide a representative basis for the derived field-structure and composition metrics.

Differences in thematic definitions among national source datasets are addressed by the TOLE reclassification step itself, which is the mechanism by which country-specific thematic schemes are harmonized into a single common legend (see ‘ALMaSS landscape map’ subsection); the specific reclassification rules applied for each input data type are documented in detail in Ziółkowska et al. (2026). Source datasets may also differ in how they were themselves originally produced by the national provider (e.g., via aerial photogrammetry, field-based survey, or farmer self-declaration with remote-sensing cross-checks); while this is not addressed in detail here, its effect on the generated landscape models and derived landscape structure metrics is considered to be of secondary importance relative to the resolution and acquisition-date

differences discussed above, given that the reclassification step operates on the final thematic output of each source dataset rather than on its production method directly.

Overview of data processing steps

ALMaSS landscape map

Details of the ALMaSS landscape generation workflow are described in Ziółkowska et al. (2026), and a brief overview is provided below. To achieve comparable detail across modeling countries/regions, despite differences in national mapping standards, the first step in ALMaSS landscape model generation was identification of suitable geodata for mapping ALMaSS TOLEs in each of the six countries selected for modeling. This produced landscape data description tables (LandDataDescriptionTable{country_code}.xlsx) and structured geodatabases or folders gathering all the input data needed for landscape model generation in each of the six countries.

The mapping algorithm (Ziółkowska et al. 2026) operates at the landscape window level, typically a 10 km × 10 km extent. Geospatial data from national or regional databases are first clipped to the landscape window extent. The ALMaSS landscape map is generated by sequentially combining individual layers of land use and land cover information with the agricultural field layer into a single raster map. Based on the landscape data description files, preprocessed layers are aggregated into thematic categories representing natural areas (Natural theme), cultivable areas (Cultivable theme, excluding agricultural fields), built-up areas (Built-up theme, excluding individual buildings), cultural features (Cultural theme), water features (Water theme), and transport infrastructure (Transport theme). Within each thematic category, layers are overlaid according to the spatial priority specified in the landscape data description files.

Single-theme raster layers and raster layers with individual buildings and agricultural fields are subsequently stacked so that ecologically meaningful layers appear at the top of the landscape map. Theme ordering depends on data characteristics and quality, prioritizing more reliable datasets, and must be adapted for each country/region. After this process, so-called 'raw' landscape maps are created. Despite careful stacking, 'raw' landscape maps usually contain unclassified 'background' areas, as well as inconsistencies arising from input data overlaps, particularly when input data vary in source or quality. To address this, a multi-stage cleaning process is then applied to remove artifacts, fill gaps, and produce a continuous landscape raster free of slivers. Large 'background' polygons with regular geometry and predominantly arable cover are converted into new agricultural fields. These represent agricultural fields not present in the current LPIS data (i.e., 'missing' fields), as not all fields may be registered for the subsidies. Elongated 'background' and sliver polygons along the fields are reclassified into field margins, while those next to waterbodies and waterlines are classified as riparian vegetation. The remaining 'background' areas, if sufficiently large and next to built-up areas, are reclassified into yards/gardens. Small 'background' objects and artifacts too minor to represent meaningful features are removed, standardizing element shapes, reducing fragmentation, and producing a more realistic landscape representation.

Finally, the cleaned maps are reclassified and regionalized to match ALMaSS landscape element types. Since these landscape maps contain more detail than required, country-specific text files guide reclassification into TOLE categories. Lastly, each feature, whether a single raster cell or multiple raster cells, receives a unique value consistent across all cells. Such a regionalization process ensures that contiguous cells representing the same feature share a common identifier. The final outputs include, for each landscape window, an ASCII raster landscape map and CSV tables containing the raster attribute table, the vector attribute table of agricultural fields, and additional properties per polygon such as soil type.

Because the availability and structure of input data varied among the mapped countries and regions, several additional preprocessing steps were required to ensure consistent landscape generation. At the level of input geodata processing, in Denmark, derivation of road and railway side slopes from the Digital Terrain Model (DTM) and Digital Surface Model (DSM) was required. In addition, nDSM data (DSM – DTM) were combined with the GeoDK hedgerow ("hegn") layer and GeoDK polygon data on buildings and forests to generate an improved vector layer of woody vegetation, including hedges, copses, and individual trees. In Finland, forest coverage was missing from the national topographic database (Maastotietokanta), so it was supplemented using canopy cover rasters from the National Forest Inventory. These rasters, representing total canopy cover and deciduous canopy cover, were used to classify five forest types: bushes (<30% canopy cover), mixed forest (>30% canopy cover and 30–70% deciduous), deciduous dominant (>30% canopy cover and >70% deciduous), coniferous dominant (>30% canopy cover and <30% deciduous), and swamp forest (>30% canopy cover overlapping the swamp land class). In Poland, where LPIS was based on cadastral parcels (Fig. 1), extensive preprocessing was required to map individual agricultural fields. This included removing non-agricultural features, dividing parcels with multiple reported crops (based on registered crop area and geometric rules), and merging adjacent parcels managed as a single field; non-agricultural objects were removed through appropriate stacking order. In Germany and the Netherlands, LPIS data included both agriculturally managed and non-managed landscape elements, so preprocessing focused on selecting arable parcels and excluding non-crop land cover types, such as forests or field margin habitats, which were treated as non-arable and grouped by land cover type. No deviations from the standard procedure were required for France.

There were also country-specific differences in the stacking of thematic layers, primarily related to the treatment of agricultural fields. Under the standard procedure, the agricultural fields layer serves as the baseline, after which the 'Cultivable' theme is added to empty space, followed by the 'Natural', 'Built-up', and 'Cultural' themes, with 'Water', 'Transport', and individual buildings added on top. In Poland, however, this order had to be modified because the reference parcels were cadastral parcels. In this case, the stacking process was also used to remove non-agricultural land-cover or land-use elements from the parcels. Therefore, agricultural parcels were stacked last, filling only the remaining empty space not covered by any other theme or by buildings.

The multi-stage cleaning process used to remove inconsistencies, artifacts, and gaps in the 'raw' landscape maps also varied across countries and re-

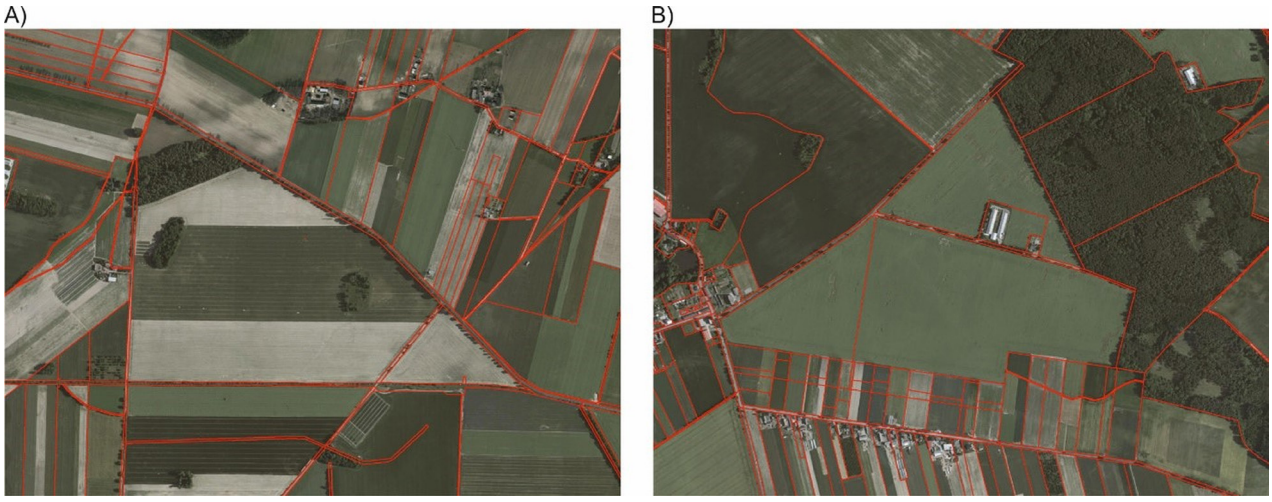

Figure 1. Examples of Polish cadastral data (red lines; EGiB WFS service, https://mapy.geoportal.gov.pl/wss/service/PZGIK/EGIB/WFS/UslugaZbiorcza) shown over the orthophoto map (GUGiK WMS service, https://mapy.geoportal.gov.pl/wss/service/PZGIK/ORTO/WMS/StandardResolution). **A.** One cadastral parcel managed as three agricultural fields; **B.** Three cadastral parcels managed as a single agricultural field.

gions. Two main issues that had to be addressed concerned grasslands and wastelands. In ALMaSS, the type and intensity of grassland management are essential for assigning polygons to the appropriate TOLE. ALMaSS distinguishes between natural (or semi-natural) grasslands (dry and wet) and managed grasslands, which can be further divided into pastures and low-yield pastures (reflecting grazing or cutting intensity), as well as amenity grass. However, this level of differentiation is not always present in the input data, as grassland classes in national topographic databases may be based solely on land cover rather than land use. In such cases, additional rules are needed to separate unmanaged from managed grasslands to avoid a substantial overestimation of natural grasslands, which are important habitats for many beneficial organisms. Alongside supplementary information layers (e.g., managed grasslands from LPIS in Poland), a set of geometry- and neighborhood-based rules was developed to reclassify general grassland polygons into more precise categories, including pastures, low-yield pastures, yards, amenity grass, field margins, and road verges. These rules are described in detail in Ziółkowska et al. (2026) using Poland as an example. Similar challenges occurred in France and the Netherlands, where grassland classes were also defined primarily by land cover, although the specific rule sets required were country-specific due to differences in landscape structure (Table 5).

A second recurring issue was the generation of excessive 'wasteland' polygons. These typically appeared as a side effect of earlier cleaning stages when background polygons or artifacts could not be assigned to any land-use category covered by the classification rules. Since wasteland objects lack defined characteristics relevant for the modeled species, they should be, where possible, reassigned to appropriate land-use classes. To address this, additional geometry- and neighborhood-based rules were created, described in detail in Ziółkowska et al. (2026). Although the general approach was applied in all countries and regions, the specific thresholds and criteria necessarily varied (Table 5).

Table 5 summarizes the general rules applied to resolve grassland and wasteland issues in five of the six countries selected for modeling. The definition of

Table 5. Geometry- and neighborhood-based rules for reassigning grassland and wasteland polygons in the landscape cleaning process, by country. Rules were applied sequentially, with each subsequent rule evaluated only for polygons not reassigned in earlier steps. All remaining general grassland polygons that did not meet any rule were assigned to Unknown grass, while unreassigned wasteland polygons remained classified as Wasteland. Denmark is not included because a different, geometry-based polygon cleaning approach was applied instead of explicit rule-based reclassification (see explanation in the text).

General rule	Germany	Finland	France	Netherlands	Poland
Area > 1 ha and regular shape	–	–	Pasture	Pasture	Pasture
Area > 0.5 ha and regular shape	–	–	Pasture	–	–
Area > 0.5 ha and irregular shape	–	–	Permanent set-aside	–	–
Area > 0.1 ha and regular shape and neighboring low built-up areas	–	–	Urban vegetation	–	–
Area > 0.1 ha and close to circular shape	–	–	Amenity grass	–	–
Area > 0.1 ha and elongated and neighboring roads	–	–	Road verge	–	–
Area > 0.1 ha and regular shape	–	–	Low-yield pasture	Low-yield pasture	Field
Area > 0.1 ha and irregular shape	–	–	Shrub		Permanent set-aside
Area > 10 m^2 and regular shape and neighboring house	–	–	Yard	Yard	Yard
Area > 5 m^2 and elongated and neighboring field	–	–	Field margin	Field margin	Field margin
Area > 5 m^2 and elongated and neighboring road	–	–	Road verge	Road verge	Road verge
Area > 0.5 ha and regular shape	–	–	–	–	Field
Irregular shape and neighboring forest	–	Forest	–	–	–
Area > 10 m^2 and regular shape and neighboring house	Yard	Yard	Yard	Yard	Yard
Area > 5 m^2 and round-like shape and inside fields	Mown grass	–	–	–	–
Area > 5 m^2 and elongated and neighboring field	Field margin	Field margin	Field margin	Field margin	Field margin
Area > 5 m^2 and elongated and neighboring road	Road verge	Road verge	Road verge	Road verge	Road verge

these rules in all countries was supported by visual inspection and verification using satellite and airplane imagery in Google Earth. In Poland, some regular, field-like grassland and wasteland polygons were additionally reclassified as fields in rotation—an adjustment not required in other countries. This occurred due to the specificity of the Polish cadastral-based LPIS data, in which many agricultural fields were absent from the register because no subsidies were claimed for them, a situation common in heterogeneous, small-scale farming landscapes. Denmark is not included in Table 5 because the input datasets already contain relatively detailed land-use information, and the cleaning procedure therefore focused mainly on removing background artifacts and assigning residual polygons based on neighboring land-use classes rather than applying explicit rule sets for grassland and wasteland reclassification.

The quality-assurance approach adopted for the ALMaSS landscape maps differs from that used for land-cover products derived through independent classification of primary source data, such as remote-sensing imagery or field survey. Because the landscape maps are compiled through semantic reclassification of pre-existing, authoritative national datasets (topographic databases, cadastral and LPIS/IACS records) whose thematic content is already quality-assured by the originating national agencies, the relevant quality dimension for this compilation step is the logical consistency of the reclassification rules, i.e., whether source classes were consistently and correctly mapped to the target TOLE categories, rather than thematic accuracy against independently collected ground-truth reference data.

After the final ALMaSS landscape maps were generated for each selected landscape window, a verification step was carried out by comparing the mapped TOLEs with imagery in Google Earth through visual inspection. Polygon-level manual adjustments, addressing residual logical-consistency errors such as implausible labels arising from neighboring context, were made where necessary; the affected proportion of each landscape window did not exceed approximately 2%, though the exact extent varied between individual landscapes.

Farming structure and cropping system

LPIS data and agricultural registers were processed in steps. First, all reference parcels were grouped into farm units based on anonymized information on reference-parcel owner/applicant identifiers or holding-unit identifiers. Second, all delineated farm units were characterized in terms of crops cultivated on fields belonging to them and in terms of animal herd/flock sizes. Each farm was assigned a specific farm type based on these characteristics.

The standard farm types considered in ALMaSS include farms specialized in field crops (arable), horticulture (vegetables and flowers), potato, beet, permanent crops (orchards—fruits, wine, and olives), grazing livestock, granivores, and mixed stock, with distinctions between conventional and organic practices. Hobby farms, defined as smallholdings that are not intended to provide a primary source of income, are also distinguished. This typology broadly corresponds to the farm categories recognized within the Farm Sustainability Data Network (FSDN; formerly FADN), although ALMaSS additionally distinguishes farms specializing in potato and beet production because these crops typically receive substantially higher levels of agronomic inputs, including fertilizers and pesticides.

To classify individual farms, country- or region-specific classification rules were developed based on crop composition and livestock characteristics. These rules were derived from analyses of Farm Structure Survey/FSDN data corresponding to the year of IACS data acquisition for each country included in the modeling (Table 6). Classification thresholds differed substantially between countries, reflecting variation in farming structure. For example, the maximum area defining hobby farms ranged from 6 ha in Poland, through 10 ha in Finland, to 20 ha in Denmark and Germany. The set of farm types also varied according to the farming systems present in each country. For example, permanent-crop farm types (orchards and vineyards) were distinguished only in France, whereas separate potato and beet farm types were defined in all other countries.

The Netherlands represented a special case, as it was the only country for which a detailed farm classification was available directly from the agricultural register. This classification, provided by the Netherlands Enterprise Agency (Rijksdienst voor Ondernemend Nederland, RVO), comprised 38 distinct farm types and therefore required simplification. Using cluster analysis based on similarities in crop acreage, the original classification was aggregated into 15 farm types, comprising animal grazing, poultry, flower, arable grazing, pig and calf, starch potato, wheat, arable vegetable, horse and sheep, fruit, greenhouse vegetable, vegetable, permanent crop, vineyard, and other farms (Ziółkowska et al. 2022).

Table 6. Farm classification rules used in the countries selected for modeling. Note that the Netherlands is not included, as farm classification for this country was provided together with the agricultural register. FA – farm area; LSU – livestock unit. Livestock unit coefficients used are based on https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Livestock_unit_(LSU).

Farm type	Denmark	Germany	Finland	France	Poland
Vegetable	> 50% vegetables in FA	(FA > 20 ha) and (> 50% vegetables in FA)	> 50% vegetables in FA	–	(FA > 6 ha) and (> 25% vegetables in FA)
Other	FA < 2 ha	–	FA < 2 ha	–	–
Potato	≥ 20% potato in FA	(FA > 20 ha) and (potato in FA ≥ 10 ha) and (≥ 20% potato in FA)	≥ 20% potato in FA	–	(FA > 10 ha) and (> 20% potato in FA)
Beet	≥ 20% beet in FA	(FA > 20 ha) and (beet in FA ≥ 16 ha) and (≥ 20% beet in FA)	≥ 20% beet in FA	–	(FA > 16 ha) and (> 20% beet in FA)
Hobby	(LSU < 20) and (FA < 20 ha)	(LSU < 20) and (FA < 20 ha)	(LSU < 20) and (FA < 10 ha)	–	(LSU < 20) and (FA < 6 ha)
Cattle	(LSU ≥ 20 and cattle + sheep > 75% in LSU) or (> 40% grazing or fodder plants in FA)	LSU ≥ 20 and cattle + sheep + goats + horse > 75% in LSU	(LSU ≥ 20 and cattle + sheep + goats > 75% in LSU) or (> 70% grazing or fodder plants in FA)	(All farms with cattle + sheep + goat herds registered) or (no herds registered but > 60% grazing or fodder plants in FA)	(LSU ≥ 20 and cattle + sheep + goats > 70% in LSU) or (> 40% grazing or fodder plants in FA)
Pig	(LSU ≥ 20) and (pigs > 75% in LSU or pig grazing areas > 0.15% of FA)	LSU ≥ 20 and pig + poultry > 26% in LSU	(LSU ≥ 20) and (pigs > 75% in LSU)	All farms with pig herds registered	(LSU ≥ 20) and (pigs > 80% in LSU)
Mixed stock	(LSU ≥ 20 and no cattle or pig farm) or (20–40% grazing or fodder plants in FA)	LSU ≥ 20 and no cattle or pig farm	(LSU ≥ 20 and no cattle or pig farm) or (20–70% grazing or fodder plants in FA)	No herds registered but 40–60% grazing or fodder plants in FA	(LSU ≥ 20 and no cattle or pig farm) or (20–40% grazing or fodder plants in FA)
Grass	–	(LSU = 0) and (≥ 40% grazing or fodder plants in FA)	–	–	–
Specialized crop	–	–	–	(FA ≥ 15 ha) and (wheat + maize + rape + barley ≥ 60% in FA)	–
Orchards	–	–	–	> 50% orchards in FA	–
Vineyards	–	–	–	> 50% vineyards in FA	–
Arable	All remaining farms	All remaining farms	All remaining farms	All remaining farms	All remaining farms

Crop selection for modeling was based on the average crop composition of each farm type. Crops occupying more than 1% of the area of a given farm type were selected as candidates for inclusion in ALMaSS. Crop management plans were then compiled for the most important crops in each country. These plans describe the sequence, timing, and probability of management operations and are used to parameterize ALMaSS crop models. The number of crop management plans implemented varied among countries (Table 7), reflecting a balance between representing crop heterogeneity and the practical effort required to collect detailed management information. In three countries, separate management plans were developed for conventional and organic production systems. Denmark had the largest number of implemented crop types, supported by additional work undertaken within the EcoMetric project. In two cases, locally sourced farmer-advisor input was not available for comparatively minor crop types within a country's cropping system, reflecting the prioritization of advisor time toward the most agriculturally important crops rather than an assessment of pedoclimatic similarity: fodder lucerne management was substituted from Poland for use in Germany (Brandenburg and Lower Saxony) and

Table 7. Types of crops (in rotation and permanent) modeled in each country (i.e., for which management plans were collected and transferred into code).

Country	No. of types of crops modeled	List of crops in rotation	List of permanent crops	If organic version available
Denmark	49	Cabbage, carrots, cereal legume, cereal legume whole, clover grass grazed, fodder beet, grazing pigs, legume beans, legume peas, legume whole, maize, maize for silage, mixed vegetables, potato, potato for industry, potato for seed, seed grass fescue, seed grass rye, set-aside, set-aside summer mow, spring barley, spring barley clover grass, spring barley for silage, spring barley green, spring fodder grass, spring oats, spring wheat, sugar beet, vegetable seeds, winter barley, winter clover grazed sown, winter fodder grass, winter rape, winter rye, winter wheat	Bush fruit, Christmas trees, energy crop, farm forest, farm young forest, grass grazed, grass grazed low yield, grass tussocky, grazing pigs, orchard apple, orchard cherry, orchard other type, orchard pear, plant nursery	Yes
Finland	21	Caraway, faba beans, feeding ground, grass grazed, grass for silage, green fallow, potato, potato for industry, starch potato, spring barley for fodder, spring barley for malt, spring oats, spring rape, spring wheat, sugar beet, turnip rape, winter rye, winter wheat	Buffer zone, green fallow, natural grassland	Yes
France	14	Grass grazed, maize, maize for silage, potatoes, sorghum, spring barley, spring oats, spring wheat, sunflower, winter barley, winter rape, winter triticale, winter wheat	Grass grazed	No
Germany	25	Cabbage, carrots, grassland silage, green fallow, herbs, legumes, maize, maize for silage, oats, peas, potatoes, potatoes for industry, spring barley, spring rye, sugar beet, triticale, winter barley, winter rape, winter rye, winter wheat	Asparagus, bush fruit, orchard, grass grazed, grass grazed low yield	Yes
Poland	14	Beans, beet, carrots, fodder lucerne, maize, maize for silage, potatoes, spring barley, spring wheat, winter barley, winter rape, winter rye, winter triticale, winter wheat	–	No
Netherlands	13	Beet, cabbage, carrots, grass grazed, grass grazed low yield, maize, potatoes, spring barley, winter wheat	Grass grazed, grass grazed low yield, orchard, tulips	No

France, and perennial herb-strip management was substituted from Germany for use in France, in both cases without parameter modification.

For each farm type in each country, a crop rotation scheme was generated based on the relative shares of crops for which crop management plans were available. These crop rotation schemes are provided to ALMaSS as *.rot* files. Each scheme consists of approximately 100 crop entries, constructed to reproduce the target long-term frequency of each crop for the given farm type. Crops are not, however, placed within the sequence according to their required frequencies alone: the sequence is constructed manually, based on agronomic knowledge and common crop-rotation practice (e.g., restricting the number of consecutive years in which particular crop groups, such as cereals, can be cultivated), and agronomic sequencing constraints are additionally enforced by ALMaSS at runtime (Topping and Duan 2024). At the start of a simulation, a random position within the rotation is assigned to each arable field, and the subsequent crop in the sequence is assumed to be grown on the same field in the following year. For example, if a crop such as maize grown for silage in an animal-grazing farm appears 13 times in a 100-entry rotation, it will, on average, occupy 13% of the fields associated with that rotation at any given time. The crop rotation represents a synthetic, agronomically plausible rotation rather than a specific field's documented cropping history, though ALMaSS also supports implementation of empirically documented, field-specific rotations where required. This was not the approach taken for the landscape datasets presented here, which are intended to represent generalized, agronomically realistic farm-type management rather than any single field's actual history.

Data characteristics

Selection of landscapes for modeling

Each country was divided into a grid of non-overlapping 10 × 10 km squares. For each square, the percentage of arable land was calculated to identify landscapes dominated by agricultural land use, where pesticide applications are a common threat to beneficial organisms. Squares with high arable land coverage were pre-selected using country-specific thresholds: >50% for Finland, >55% for France, and >60% for the remaining countries, as it was not possible to identify sufficient numbers of 10 × 10 km squares with ≥60% arable land in Finland and France.

For all pre-selected squares, the total number of fields was derived and the mean field size calculated. From this pool, 10 landscape windows per country were manually selected to represent the widest possible gradient in landscape structure, particularly in terms of the number of fields and the mean field size (Fig. 2). It should be noted that these selected landscape windows are not representative of national averages; they were deliberately chosen to capture agricultural heterogeneity and therefore reflect conditions in agriculturally active areas rather than whole-country statistics.

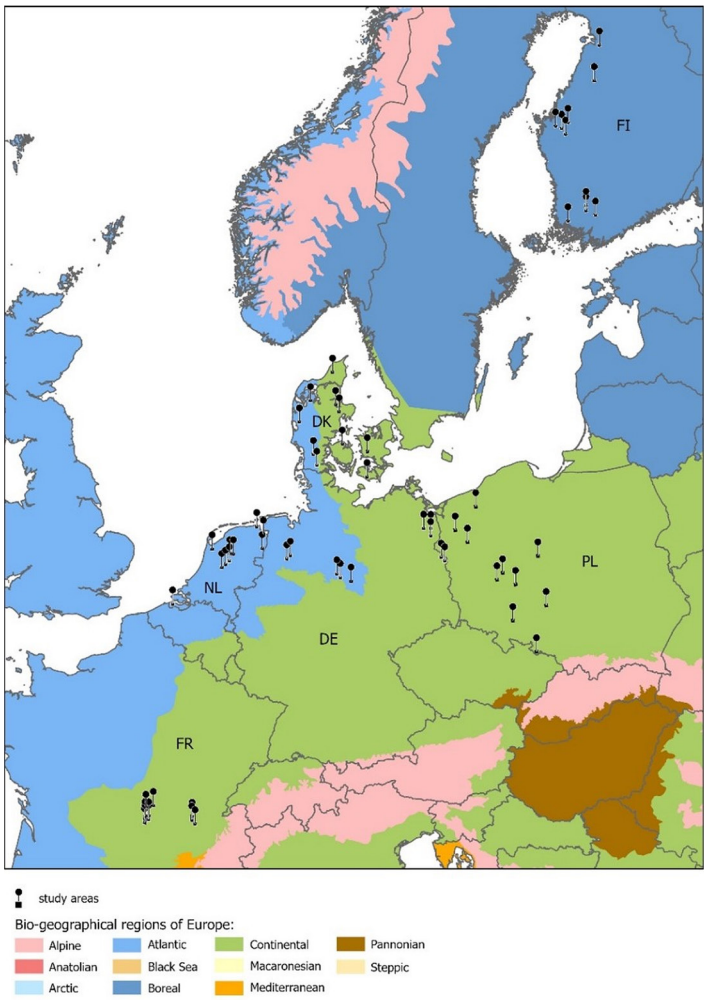

Figure 2. Location of landscape windows (10 × 10 km) selected for modeling. Source of the map of biogeographical regions of Europe: European Environment Agency (EEA 2002).

Landscape structure of selected landscape windows

Selected landscape windows were characterized by different landscape metrics covering four aspects of agricultural landscape structure: land cover composition, field and farm structure, farm-type diversity, and semi-natural habitat structure and connectivity. Table 8 reports the full set of these metrics; metrics disaggregated by semi-natural habitat type (grassy vs. woody) and landscape/farm configuration indices are provided in Suppl. material 1: table S2, and full formulas and definitions for every metric are provided in Suppl. material 1: table S3. For patch-based metrics (e.g., semi-natural habitat patch counts, areas, and inter-patch distances), a patch is defined as a maximal set of spatially contiguous, same-class raster cells under an 8-neighbor (queen's case) connectivity rule, with no minimum patch size applied; edge-length and cell-adjacency-based metrics (e.g., arable–semi-natural boundary length) instead use 4-neighbor cell-side adjacency, the standard convention in FRAGSTATS-equivalent software, since patch delineation and edge/adjacency counting intentionally use different neighborhood rules.

To summarize the main axes of landscape variation using a non-redundant set of variables suitable for multivariate analysis, a principal component analysis (PCA) was performed on a subset of 16 of these metrics, spanning land cover composition, landscape and farm-level configuration, and field structure (marked † in Table 8 and Suppl. material 1: table S2). This subset excludes metrics disaggregated by habitat type or otherwise representing finer-grained detail than needed to characterize whole-landscape structure (e.g., grassy- and woody-specific patch statistics), since including multiple disaggregated versions of closely related metrics would reintroduce the kind of redundancy a non-redundant variable set is intended to avoid. Prior to PCA, each metric's skewness was assessed and, where warranted, a transform applied; metrics were then standardized. A hierarchical cluster analysis (Ward linkage) was performed on the resulting principal component scores; the clustering input (PC1–PC2 scores) and number of clusters were both determined empirically by comparing internal validity diagnostics across candidate options, rather than assumed. Full details of the transformation, correlation structure, and clustering-comparison procedures are provided in Suppl. material 1: S4.

Landscape windows differed considerably in all these dimensions. Finnish landscape windows were characterized by very high woody cover (mean 28.8 ± 5.9%) (Fig. 3), reflecting the forest matrix in which Finnish agricultural land is embedded, whereas landscape windows in the Netherlands featured higher coverage of semi-natural grassy elements and water features (Table 8). The largest-scale farming was found in landscape windows located in the Brandenburg region in Germany (mean field size between 9.3 and 12.2 ha). Medium-sized fields dominated in Danish (mean field size between 4.5 and 9.0 ha) and Dutch (mean field size between 3.4 and 6.9 ha) landscape windows. Landscape windows in the Lower Saxony region in Germany, France, and Finland were dominated by smaller fields (2.5–5.0 ha, 2.2–3.3 ha, and 2.4–4.4 ha, respectively) (Fig. 4; Table 8). The greatest within-country variation occurred in Poland, where mean field sizes ranged from 0.7 to 3.9 ha (Fig. 4) and the number of arable fields per landscape window varied from approximately 2,100 to 11,700. Farm-type diversity was highest in landscape windows in Denmark and

Table 8. Summary of landscape metrics in selected agricultural regions across six European countries. Values are mean ± standard deviation across 10 landscapes per country (*n* = 10). DE = Germany, DK = Denmark, FI = Finland, FR = France, NL = Netherlands, PL = Poland. SNH = semi-natural habitat. Areas originally computed from raster data (m^2) are reported in hectares (ha). NN = nearest neighbor. Landscape- and farm-configuration metrics (LSI, IJI, COHESION, DIVISION, SPLIT, SIDI, and their farm-level equivalents) and grassy/woody-specific patch statistics are provided in Suppl. material 1: table S2. Full formulas and definitions for every metric are provided in Suppl. material 1: table S3.

Metric	Description	Unit	DE	DK	FI	FR	NL	PL
Land cover composition								
Arable cover	Proportion of landscape covered by all arable types	%	80.30 ± 4.52	78.51 ± 3.43	59.93 ± 6.04	73.29 ± 6.59	76.78 ± 3.34	82.54 ± 6.48
Grassy cover	Proportion covered by semi-natural grassy habitats	%	4.61 ± 1.07	4.38 ± 0.94	4.25 ± 1.62	3.64 ± 1.11	6.17 ± 1.64	2.72 ± 1.05
Woody cover	Proportion covered by woody/forest habitats	%	7.13 ± 2.95	7.81 ± 3.89	28.81 ± 5.86	11.26 ± 5.73	3.82 ± 2.66	9.06 ± 6.60
Built cover	Proportion covered by built-up/impervious surfaces	%	6.94 ± 2.82	8.42 ± 1.75	5.23 ± 1.65	9.42 ± 4.56	9.04 ± 3.60	4.95 ± 1.61
Water cover	Proportion covered by water bodies	%	1.01 ± 0.62	0.88 ± 0.42	1.78 ± 0.61	2.39 ± 2.23	4.17 ± 2.89	0.73 ± 0.97
Arable land use								
Rotational crops †	Proportion of landscape under rotational crops	%	76.14 ± 6.18	74.20 ± 4.24	59.26 ± 5.77	64.54 ± 6.35	67.83 ± 2.76	76.13 ± 8.62
Permanent grassland †	Proportion under permanent grassland/pasture	%	4.01 ± 2.41	4.07 ± 1.36	0.64 ± 0.37	8.58 ± 3.80	5.14 ± 3.30	5.71 ± 3.75
Fruit and nut production	Proportion under orchards/ fruit and nut production	%	0.11 ± 0.17	0.08 ± 0.12	0.03 ± 0.03	0.14 ± 0.11	1.52 ± 1.93	0.69 ± 0.74
Semi-natural composition								
Herbaceous SNH cover †	Proportion covered by semi-natural herbaceous habitats	%	3.43 ± 1.11	2.74 ± 0.60	3.85 ± 1.18	3.41 ± 1.08	5.08 ± 1.29	2.52 ± 1.00
Woody SNH cover †	Proportion covered by semi-natural woody habitats	%	7.13 ± 2.95	7.24 ± 3.38	28.80 ± 5.86	11.10 ± 5.69	3.73 ± 2.64	9.06 ± 6.60
Field and farm structure								
No. arable fields †	Count of arable field polygons per landscape	n	1508 ± 845	1467 ± 247	1849 ± 333	2840 ± 492	1785 ± 382	5989 ± 2878
Mean field size †	Mean area of arable field polygons	ha	7.26 ± 3.96	5.58 ± 1.31	3.36 ± 0.76	2.64 ± 0.38	4.53 ± 1.23	1.76 ± 1.00
Mean farm area	Mean total arable area per farm holding	ha	93.53 ± 63.85	62.59 ± 15.27	33.77 ± 6.45	36.68 ± 11.09	37.05 ± 11.35	20.91 ± 11.75
Gini (farm size)*	Inequality in farm size distribution	0–1	0.66 ± 0.10	0.66 ± 0.03	0.52 ± 0.02	0.59 ± 0.01	0.52 ± 0.04	0.67 ± 0.10
Farm-type diversity								
Shannon div. (fields)**	Shannon diversity of farm types weighted by field count	nats	1.28 ± 0.17	1.58 ± 0.18	1.27 ± 0.22	0.45 ± 0.18	1.11 ± 0.18	0.50 ± 0.28
Shannon div. (area)** †	Shannon diversity of farm types weighted by arable area	nats	1.15 ± 0.14	1.46 ± 0.20	1.23 ± 0.22	0.40 ± 0.17	1.08 ± 0.25	0.71 ± 0.31
Semi-natural habitat patches								
No. SN patches	Count of all SN patches (grassy + woody combined)	n	1638 ± 380	14221 ± 2893	1099 ± 453	5377 ± 1206	1697 ± 628	2630 ± 1401
Mean SN patch area	Mean area of combined SN patches	ha	0.66 ± 0.22	0.07 ± 0.03	3.36 ± 1.22	0.30 ± 0.19	0.56 ± 0.26	0.60 ± 0.58
SN NN distance (centroid) ***	Centroid-to-centroid distance between nearest SN patches – spacing	m	75.1 ± 13.1	20.7 ± 3.3	111.7 ± 23.6	48.0 ± 6.7	78.6 ± 13.1	78.0 ± 22.5
SN NN distance (edge-to-edge) ***	Edge-to-edge distance between nearest SN patches – connectivity	m	13.9 ± 2.7	5.0 ± 0.4	14.6 ± 2.3	7.3 ± 1.0	9.1 ± 1.1	14.4 ± 4.6

Metric	Description	Unit	DE	DK	FI	FR	NL	PL
Arable–semi-natural edge metrics								
Arable-SN edge length	Total length of arable/semi-natural boundary	km	1135.7 ± 364.8	1109.3 ± 242.5	1120.8 ± 95.5	1496.7 ± 145.4	1269.7 ± 264.8	1023.6 ± 242.7
Arable-SN edge density	Arable-SN edge length per hectare of arable land	m/ha	143.9 ± 55.0	141.4 ± 30.3	189.1 ± 28.6	204.5 ± 13.4	166.2 ± 38.5	124.3 ± 30.3

* Gini coefficient measures inequality in farm size distribution, ranging from 0 (all farms equal in size) to 1 (one farm holds all the area).
** Shannon diversity index quantifies both the number of farm types present and the evenness of their representation—a landscape with many equally common farm types has a higher value than one dominated by a single type. Values are expressed in nats (natural units of information obtained when the natural logarithm is used); higher values indicate greater farm-type diversity.
*** Centroid-to-centroid distance measures inter-patch spacing. Edge-to-edge distance approximates inter-patch connectivity. Mean nearest-neighbor distance reflects local, pairwise spacing/connectivity and does not capture landscape-wide patch clustering; COHESION (land) (Suppl. material 1: table S2) provides a whole-landscape measure of physical connectedness accounting for the full patch configuration.
† Metric used in the 16-variable PCA/landscape typology (Fig. 6), together with others presented in Suppl. material 1: table S2.

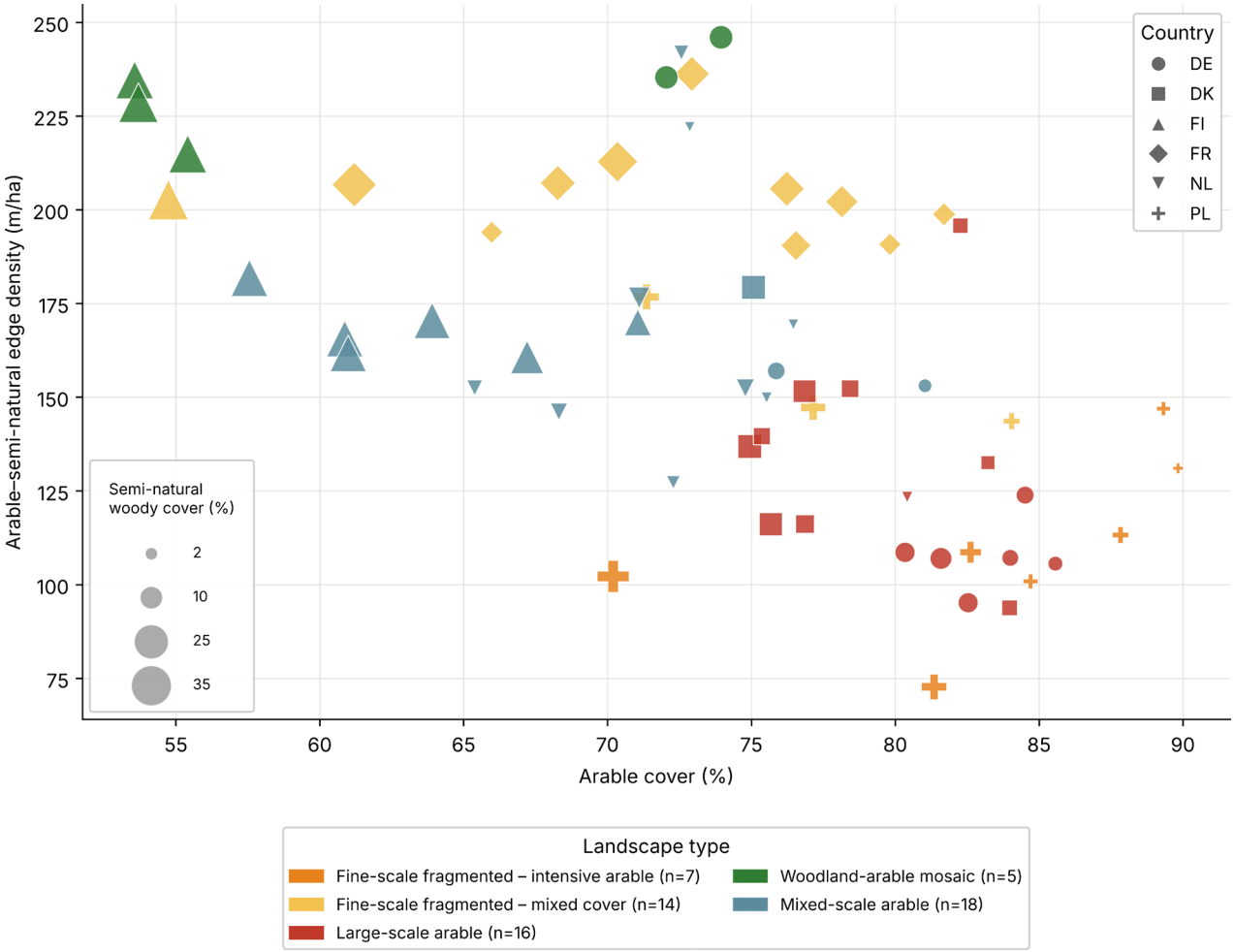

Figure 3. Relationship between arable land cover and arable–semi-natural edge density across the 60 landscape windows. Symbol size represents woody vegetation cover. Symbol shape indicates country, and color indicates landscape cluster (type) assigned by hierarchical cluster analysis. DE = Germany, DK = Denmark, FI = Finland, FR = France, NL = Netherlands, PL = Poland.

lowest in France and Poland, where one or two farm types dominated (Fig. 4). Danish landscape windows contained by far the highest number of semi-natural patches, with correspondingly the smallest mean patch area, forming an extremely fine-grained network; French and, to a lesser extent, Polish landscape

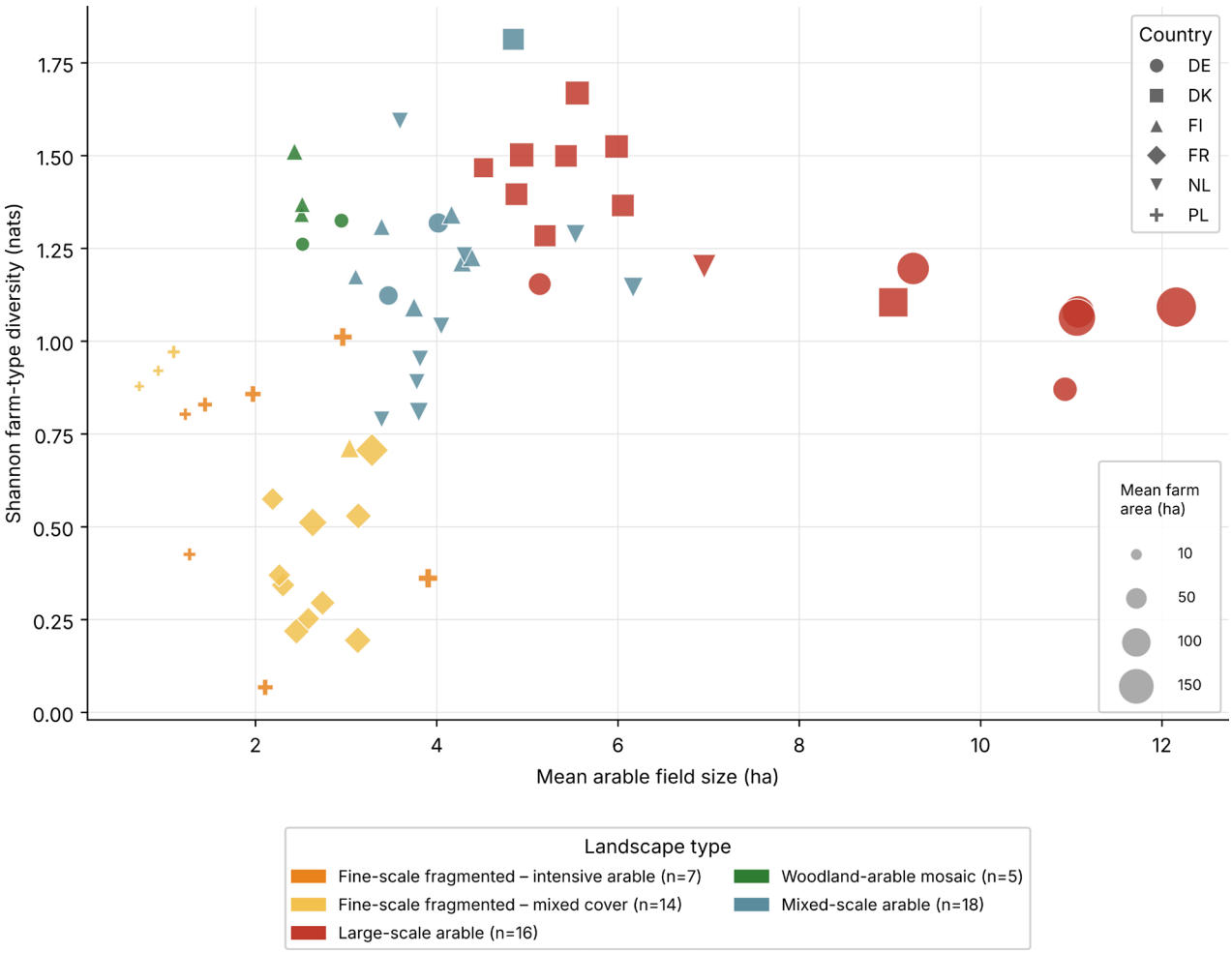

Figure 4. Relationship between mean arable field size and Shannon farm-type diversity across the 60 landscape windows. Symbol size represents mean farm area. Symbol shape indicates country, and color indicates landscape cluster (type) assigned by hierarchical cluster analysis. DE = Germany, DK = Denmark, FI = Finland, FR = France, NL = Netherlands, PL = Poland.

windows also showed relatively high patch counts and small mean areas, while Finnish landscape windows hosted fewer but substantially larger patches (Fig. 5). Arable–semi-natural edge density (the total length of edges between arable and semi-natural habitats per area of arable land) was highest in landscape windows in France and lowest in Poland, with the remaining four countries occupying an intermediate range (Fig. 3).

The first two principal components (PC1 and PC2) explained 28.1% and 19.1% of total variance, respectively (47.2% combined). PC1 separated landscapes dominated by arable land with comparatively simple, homogeneous composition from landscapes with greater configurational complexity: higher semi-natural habitat cover (both herbaceous and woody), greater landscape and farm-level fragmentation and diversity, and more convoluted patch shapes. PC2 represented a gradient of farming scale and diversity, separating landscapes with large fields and high farm-type diversity from fine-grained systems with small, arable-dominated parcels and high landscape cohesion. Based on the PCA scores, a hierarchical cluster analysis identified five landscape types (mean standardized values of all 16 metrics per cluster are shown in Suppl. material 1: fig. S4.5). The fine-scale fragmented–intensive arable clus-

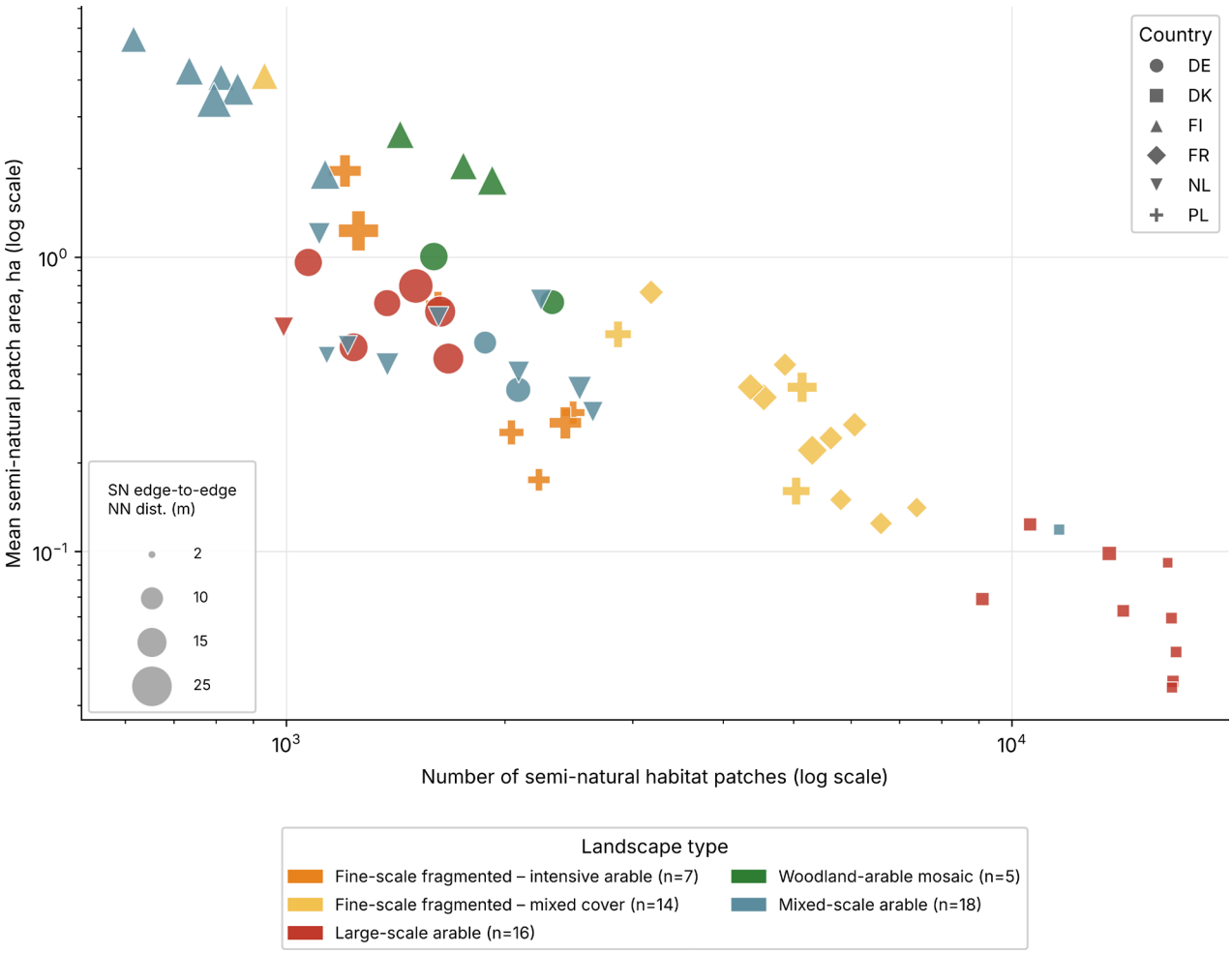

Figure 5. Relationship between the number of semi-natural habitat patches and mean patch area across the 60 landscape windows. Both axes are on a log scale. Symbol size represents the mean edge-to-edge nearest-neighbor distance between semi-natural patches, a measure of inter-patch connectivity (see Table 8 for the corresponding centroid-to-centroid spacing metric). Symbol shape indicates country, and color indicates landscape cluster (type) assigned by hierarchical cluster analysis. DE = Germany, DK = Denmark, FI = Finland, FR = France, NL = Netherlands, PL = Poland.

ter (*n* = 7) comprised exclusively Polish landscape windows, characterized by the smallest mean field size, the highest number of arable fields, and the lowest semi-natural herbaceous cover of any cluster. The fine-scale fragmented–mixed cover cluster (*n* = 14) grouped most French landscape windows together with a minority of Polish and one Finnish window, sharing small fields and high field counts with the previous cluster but with lower arable dominance and greater woody and herbaceous cover. The large-scale arable cluster (*n* = 16) contained the most consolidated landscape windows, drawn predominantly from Denmark and Germany with one Dutch window, and was characterized by the largest fields and the largest farm holdings. The woodland-arable mosaic cluster (*n* = 5) comprised a minority of Finnish and German landscape windows with the highest woody cover, the lowest arable cover, and the most fragmented configuration of any cluster. The mixed-scale arable cluster (*n* = 18) was the most compositionally diverse, comprising the majority of Dutch and Finnish landscape windows together with a minority of German and Danish ones, and was characterized by medium field sizes and the highest herbaceous semi-natural cover. Together, the five clusters span a broad range of structural land-

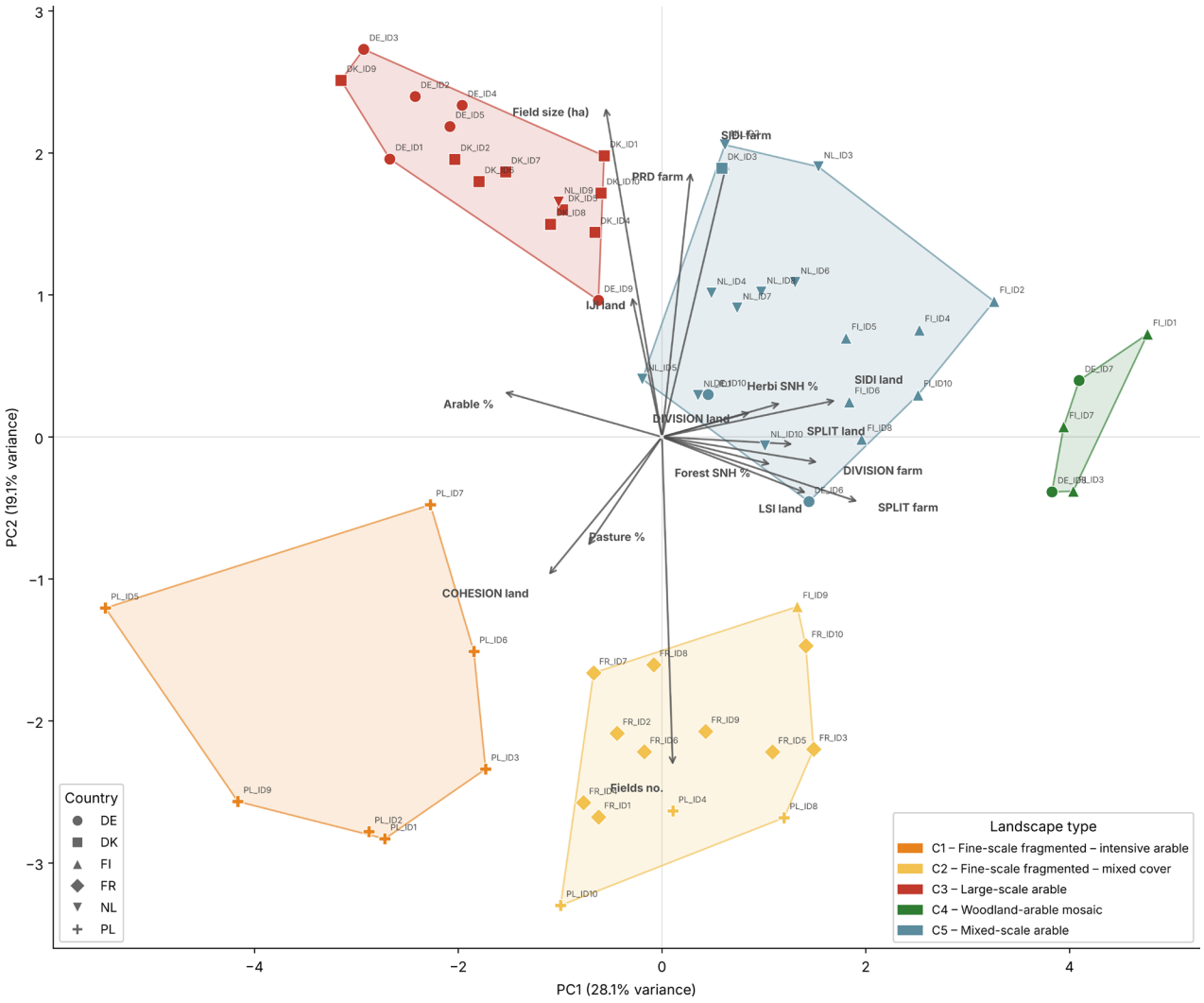

Figure 6. Principal component analysis of landscape structure metrics across 60 European agricultural landscape windows. Each point represents one landscape window; symbol shape indicates country, and symbol color indicates landscape cluster (type) assigned by Ward hierarchical cluster analysis. Arrows show loadings of the 16 landscape metrics on the first two principal components. Shaded convex hulls delineate the five landscape clusters (types): C1 – Fine-scale fragmented–intensive arable (orange; Polish landscapes with the smallest fields, the highest field counts, and the lowest semi-natural herbaceous cover); C2 – Fine-scale fragmented–mixed cover (yellow; predominantly French landscapes with small fields and higher woody and herbaceous cover than C1); C3 – Large-scale arable (red; predominantly Danish and German landscapes with large fields and farm holdings and the highest farm-size inequality); C4 – Woodland-arable mosaic (green; Finnish and German landscapes with the highest woody cover and the most fragmented configuration); C5 – Mixed-scale arable (teal; predominantly Dutch and Finnish landscapes with medium field sizes and the highest herbaceous semi-natural cover). PC1 and PC2 explain 28.1% and 19.1% of total variance, respectively. DE = Germany, DK = Denmark, FI = Finland, FR = France, NL = Netherlands, PL = Poland.

scape types, demonstrating that the landscape windows selected for modeling represent a substantial diversity of European agricultural landscapes in terms of land use composition, farming scale, and semi-natural habitat configuration.

Data resources

All data are permanently archived and available via Zenodo at https://doi.org/10.5281/zenodo.22935170 (release v. 1.0.0, archived 24 September 2026),

with checksums provided for each distributed file. The underlying repository is maintained in a GitLab repository (https://gitlab.com/ALMaSS/almass_methodology/-/tree/main/landscape_examples) for ongoing updates and development; the archived Zenodo release corresponds to the version of the repository described in this paper. All 60 landscape models have been used operationally within the EcoStack project, where they were run with multiple ALMaSS species models to produce simulation outputs (Chetcuti et al. 2024), confirming that they load and execute correctly.

The repository contains ALMaSS-ready input files for 60 landscape models, organized hierarchically by country and file type. At the top level, the repository contains six country folders (one per country, named by the two-letter country code {CC}, e.g., PL for Poland), a Common_files folder holding shared configuration files used across all countries, and a Script folder containing a Jupyter notebook for generating visualizations and GIS files from the landscape input files.

Each country folder contains two country-level files alongside three subfolders. The landscape data description table (LandDataDescriptionTable{CC}.xlsx) documents input datasets, processing rules, and the floral resource curves used in the simulations. The farm type lookup file ({CC}_FarmType_lookup.txt) defines the mapping between numeric farm type codes used in the {CC}_FarmRef.txt files and descriptive farm type names. The Landscape_files subfolder contains the core ALMaSS input files for all 10 landscape windows: {CC}_ID{X}.lsb, {CC}_ID{X}_PolyRef.txt, and {CC}_ID{X}_FarmRef.txt, where {X} denotes the landscape identifier (1 to 10), as described in Table 1. The Rotations subfolder stores crop rotation scheme files (.rot) covering all farm types defined for the given country. The Weather_files subfolder contains weather input files (.pre) for each of the 10 landscape windows.

Raster landscape maps in GeoTIFF (.tif) format, enabling direct use in GIS software, can be generated from the ALMaSS landscape input files using the Jupyter notebook provided in the Script folder. The notebook reads the {CC}_ID{X}.lsb, {CC}_ID{X}_PolyRef.txt, and {CC}_ID{X}_FarmRef.txt files for a specified country and landscape window and produces three thematic maps—polygon TOLE type, farm type, and farm objects—as both figure exports and georeferenced GeoTIFF files.

The Common_files folder contains shared configuration files used across all countries, including vegetation growth curve definitions and polygon type class definitions. Where available, the country folder also contains documentation files describing the landscape generation workflow. Detailed descriptions of the landscape generation process are provided for Finland, Denmark, and France within the repository, while equivalent descriptions for Poland, the Netherlands, and Germany are available in published sources (Poland: Ziółkowska et al. 2021; the Netherlands: Ziółkowska et al. 2022; Germany: Ziółkowska et al. 2024). Sources for all input datasets used to generate the ALMaSS landscape models for the six selected countries are described in earlier sections of this paper.

Usage licence

The set of ALMaSS-ready input files for the 60 landscapes described in this paper is released under the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits use, sharing, adaptation, and redistribution

in any medium or format, provided appropriate credit is given to the original authors and source.

The datasets represent derived products generated for use in the ALMaSS modeling framework. While the processed landscape models are freely available under CC BY 4.0, some underlying input datasets (e.g., agricultural registers) are subject to national or institutional access restrictions and are not redistributed in their original form. Users are responsible for complying with the licensing conditions of any third-party data if they reconstruct, modify, or update the landscape models.

The datasets are provided without warranty of any kind. The authors and data providers accept no responsibility for any use made of the data.

Users of the datasets are requested to cite both this data paper and the relevant ALMaSS framework publications when using the landscape models in scientific or applied work.

Additional information

The complete workflow for generating landscape models for ALMaSS was implemented in a set of Python scripts using the Pandas library (https://pandas.pydata.org/). These scripts were adapted separately for each country or region to account for differences in available spatial data and data structures. The workflow consists of three main scripts.

The first script, *{country_code}_landscape_generation_part1.py*, performs initial preprocessing of input geodata, including clipping to the landscape window, querying and converting datasets, combining individual raster layers into thematic maps, and stacking these themes to produce a preliminary "raw" landscape map. The second script, *{country_code}_landscape_generation_part2.py*, applies a multi-stage procedure to resolve inconsistencies arising from data artifacts and gaps in the raw map. This script also performs reclassification and regionalization and exports the resulting outputs for further processing. These outputs include the final polygon-based landscape raster map in ASCII format, the corresponding raster attribute table in .csv format, the attribute table of the vector layer representing agricultural fields in .csv format, and optional .csv tables containing additional attributes of landscape polygons.

The third script, *{country_code}_ALMaSS_input_files_generation.py*, uses the outputs from the previous step to generate the landscape input files required by ALMaSS. In addition to the processed landscape data, this script requires access to the agricultural register in order to assign agricultural fields in the landscape raster map to individual farm units. Farm classification is carried out using an additional script, *{country_code}_farm_classification.py*, which applies country-specific farm classification rules to assign farm types; the resulting classification forms an input to *{country_code}_ALMaSS_input_files_generation.py*.

The landscape generation scripts are described in detail in Ziółkowska et al. (2026), using an example from Poland. The scripts rely on Python and the ArcPy site package to access ArcGIS functionality and can be executed within the ArcGIS environment. The versions of the scripts used to generate the landscape models presented in this paper were executed using Python 2.7 and ArcPy within ArcGIS 10.x. These versions are available from the authors upon request; however, depositing them in full would not, in any case, be sufficient to

make the complete landscape-generation workflow independently reproducible, since some of the underlying source data, including parts of the national IACS databases, are subject to access or redistribution restrictions and cannot be archived or shared alongside the scripts. To support reproducibility and, more directly, reuse of the resulting landscapes, the complete set of ALMaSS-ready input files generated by this workflow for each country is instead provided (see Data resources), which documents and enables direct use of the published landscapes independently of the restricted raw source data. The scripts are currently being reconstructed for Python 3.x and updated to accommodate newer input datasets as part of the Horizon 2020 PollinERA project. In addition, efforts are underway to develop an open-source implementation of ALMaSS landscape generation utilizing Python packages from the osgeo family, specifically GDAL and OGR.

Acknowledgments

This work was supported by the EcoStack project funded by the European Union's Horizon 2020 research and innovation Programme under Grant Agreement No. 773554, the PBL Netherlands Environmental Assessment Agency through the project "Developing and application of a methodology to assess impacts of pesticides on key ecosystem services", and the UBA German Environment Agency through the ELONTA project "Evaluating the impact of landscape structure and source-sink dynamics on non-target arthropod pesticide risk assessments in Germany".

References

Chetcuti J, Ziółkowska E, Andersen AH, Thomsen P, Duan X, Topping CJ (2024) Results of integrative analyses – European scale analysis of impacts of EcoStack innovations and suggested management strategies (Northern) (Deliverable No. D8.3). EcoStack, Horizon 2020 Research and Innovation Programme under Grant Agreement No. 773554.

EEA (2002) Europe's biodiversity – biogeographical regions and seas (EEA Report No. 1/2002). European Environment Agency. https://www.eea.europa.eu/en/analysis/publications/report_2002_0524_154909

Fahrig L, Baudry J, Brotons L, Burel FG, Crist TO, Fuller RJ, Sirami C, Siriwardena GM, Martin J-L (2010) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. Ecology Letters 14: 101–112. https://doi.org/10.1111/j.1461-0248.2010.01559.x

Gamma E, Helm R, Johnson R, Vlissides J (1994) Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, 195 pp. https://doi.org/10.1007/3-540-47910-4_21

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146: 1999–2049. https://doi.org/10.1002/qj.3803

Hokkanen HMT (2017) Ecostacking: maximising the benefits of ecosystem services. Arthropod-Plant Interactions 11: 741–742. https://doi.org/10.1007/s11829-017-9575-8

Martínez-Núñez C, Kleijn D, Ganuza C, Heupink D, Raemakers I, Vertommen W, Fijen TPM (2022) Temporal and spatial heterogeneity of semi-natural habitat, but not crop diversity, is correlated with landscape pollinator richness. Journal of Applied Ecology 59: 1258–1267. https://doi.org/10.1111/1365-2664.14137

Priyadarshana TS, Martin EA, Sirami C, Woodcock BA, Goodale E, Martínez-Núñez C, Lee M-B, Pagani-Núñez E, Raderschall CA, Brotons L, Rege A, Ouin A, Tscharntke T, Slade EM (2024) Crop and landscape heterogeneity increase biodiversity in agricultural landscapes: A global review and meta-analysis. Ecology Letter 27: e14412. https://doi.org/10.1111/ele.14412

Topping CJ (2022) The Animal Landscape and Man Simulation System (ALMaSS): A history, design, and philosophy. Research Ideas and Outcomes 8: e89919. https://doi.org/10.3897/rio.8.e89919

Topping CJ, Duan X (2024) ALMaSS Landscape and Farming Simulation: Software classes and methods. Food and ecological Systems Modelling Journal 5: e121215. https://doi.org/10.3897/fmj.5.121215

Topping CJ, Hansen T (2003) ALMaSS, an agent-based model for animals in temperate European landscapes. Ecological Modelling 167: 65–82. https://doi.org/10.1016/S0304-3800(03)00173-X

Tscharntke T, Klein AM, Kruess A, Steffan-Dewenter I, Thies C (2005) Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. Ecology Letters 8: 857–874. https://doi.org/10.1111/j.1461-0248.2005.00782.x

Ziółkowska E, Topping CJ, Bednarska AJ, Laskowski R (2021) Supporting non-target arthropods in agroecosystems: Modelling effects of insecticides and landscape structure on carabids in agricultural landscapes. Science of the Total Environment 774: e145746. https://doi.org/10.1016/j.scitotenv.2021.145746

Ziółkowska E, Tiktak A, Topping CJ (2022) Is the effectiveness of policy-driven mitigation measures on carabid populations driven by landscape and farmland heterogeneity? Applying a modelling approach in the Dutch agroecosystems. PLOS ONE 17(12): e0279639. https://doi.org/10.1371/journal.pone.0279639

Ziółkowska E, Laskowski R, Dominic AR, Stein S, Wenzel B, Kehlenbeck H, Riddervold M, Topping CJ (2024) Evaluating the impact of landscape structure and source-sink dynamics on non-target arthropod pesticide risk assessments in Germany. German Environment Agency, Texte | 58/2024, 150 pp. https://doi.org/10.60810/openumwelt-6012

Ziółkowska E, Jaśkowiec B, Groom GB, Topping CJ (2026) Development of a spatio-temporal representation of agricultural landscapes as the modelling environment for spatially explicit agent-based models in the Animal Landscape and Man Simulation System (ALMaSS). Agricultural and Environmental Modelling 8: e167439. https://doi.org/10.3897/aem.8.167439

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Artificial Intelligence (AI) use

The authors accept full responsibility for the content of the manuscript, including the disclosure of any use of AI.
Claude (Anthropic) was used for language editing of the authors' text. Claude was also used as a coding assistant in developing the scripts for landscape metrics calculation and data analysis. All AI-assisted outputs were reviewed, edited and verified by the authors, who take full responsibility for the final text, code and results.

Funding

EcoStack project funded by the European Union's Horizon 2020 research and innovation Programme under Grant Agreement No. 773554; UBA German Environment Agency through the ELONTA project "Evaluating the impact of landscape structure and source-sink dynamics on non-target arthropod pesticide risk assessments in Germany"; PBL Netherlands Environmental Assessment Agency through the project "Developing and application of a methodology to assess impacts of pesticides on key ecosystem services".

Author contributions

Elżbieta Ziółkowska: Conceptualization, data curation (Netherlands, Poland), resources (Netherlands, Poland), formal analysis (Netherlands, Poland), software (Germany, Finland, France, Netherlands, Poland), supervision, validation, visualization, writing—original draft. Geoff Brian Groom: Data curation (Denmark), resources (Denmark), formal analysis (Denmark), software (Denmark), validation, writing—review and editing. Jean-Sébastien Beaulne: Data curation (France), resources (France), investigation (France), formal analysis (France), software (France), writing—review and editing. Anto Raja Dominic: Resources (Germany), writing—review and editing. Hella Kehlenbeck: Resources (Germany), writing—review and editing. Luna Kondrup Marcussen: Resources (Denmark), investigation (Denmark), software (Denmark), writing—review and editing. Łukasz Mikołajczyk: Data curation (Finland), formal analysis (Finland), software (Finland), writing—review and editing. Anne Muola: Resources (Finland), writing—review and editing. Sari Peltonen: Resources (Finland), investigation (Finland), writing—review and editing. Benoit Sarrazin: Resources (France), software (France), writing—review and editing. Jürgen Schwarz: Resources (Germany), writing—review and editing. Susanne Stein: Data curation (Germany), resources (Germany), formal analysis (Germany), software (Germany), writing—review and editing. Bettina Wenzel: Resources (Germany), investigation (Germany), writing—review and editing. Christopher John Topping: Funding acquisition, software, supervision, writing—review and editing.

Author ORCIDs

E. Ziółkowska https://orcid.org/0000-0002-7213-2200
G. Groom https://orcid.org/0000-0003-3820-6621
A.R. Dominic https://orcid.org/0000-0003-0754-6199
H. Kehlenbeck https://orcid.org/0000-0002-5568-4070
L.K. Marcussen https://orcid.org/0000-0002-3176-2281
Ł. Mikołajczyk https://orcid.org/0000-0002-5593-9854
A. Muola https://orcid.org/0000-0003-1828-6425
B. Sarrazin https://orcid.org/0000-0001-7310-321X
J. Schwarz https://orcid.org/0000-0002-8678-367X
B. Wenzel https://orcid.org/0000-0002-4885-5062
C.J. Topping https://orcid.org/0000-0003-0874-7603

Data availability

All data that support the findings of this study are available in the main text, in the Supplementary Information, or in the Zenodo repository at https://doi.org/10.5281/zenodo.22935170 (version 1.0.0, archived 24 September 2026).

Supplementary material 1

Supplementary information

Authors: Elżbieta Ziółkowska, Geoff Groom, Jean-Sébastien Beaulne, Anto Raja Dominic, Hella Kehlenbeck, Luna Kondrup Marcussen, Łukasz Mikołajczyk, Anne Muola, Sari Peltonen, Benoit Sarrazin, Jürgen Schwarz, Susanne Stein, Bettina Wenzel, Christopher John Topping

Data type: docx

Explanation note: **table S1.** List and characteristics of Types of Landscape Elements (TOLEs) used in the generated ALMaSS landscape maps. **table S2.** Additional landscape- and farm-configuration metrics. **table S3.** Full metric definitions. S4: PCA and Clustering Methodology. **fig.S4.1.** PCA diagnostics. **S4.1.** Variable transformation. **S4.2.** Correlation structure and redundancy. **fig. S4.2.** Pairwise Pearson correlation matrix of the 16 PCA metrics (post-transformation, pre-standardization). **S4.3.** Clustering input comparison. **fig. S4.3.** Cluster-validity diagnostics (silhouette, Davies-Bouldin, and Calinski-Harabasz) across K = 2–10, compared across the three candidate clustering inputs (rows). **S4.4.** Cluster number selection. **fig. S4.4.** Ward-linkage dendrogram on PC1–PC2 scores (cophenetic correlation = 0.670). **fig. S4.5.** Cluster profiles: mean standardized (z-score) value of each of the 16 PCA metrics per cluster, supporting the cluster descriptions in the main text (Results). **S4.5.** Reproducibility.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/aem.8.205420.suppl1

Supplementary material 2

Supplementary code (with zip folder "Supplementary code for landscape analysis")

Author: Elżbieta Ziółkowska

Data type: zip

Explanation note: landscape_metrics.csv - landscape metrics calculated for all 60 landscape map. landscape_metrics.ipynb - Jupyter Notebook script used to calculate landscape metrics based on ALMaSS landscape maps. pca_ward_clustering.ipybn - Jupyter Notebook script used for Principal Component Analysis (PCA) on landscape metrics and landscape clustering.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/aem.8.205420.suppl2