ETC DI Report 2026/6

Summary report on potential methodological approach and resource needs for a regular EU ecosystem extent data product

Authors:
Emanuele Mancosu, Wouter Plouvier, Antonio Sanchez (UMA), Roger Milego (UAB), Jaroslav Dufek (GISAT), Stefan Kleeschulte (s4e), Gergely Maucha, Barbara Kosztra (Lechner), Pavel Milenov, Jan-Erik Petersen (EEA)

European Environment Agency
European Topic Centre
Data integration and digitalisation

Cover design: EEA
Cover image © Emanuele Mancosu
Layout: EEA, ETC DI, UMA

Version: 1.0

Publication Date: 28 August 2026

EEA activity: Biodiversity and Ecosystems

ETC DI coordinator: Umweltbundesamt GmbH (UBA) / Environment Agency Austria (EAA)

ETC DI partners: University of Malaga, Space4Environment Sarl, Universitat Autonoma de Barcelona, GISAT S.R.O., Epsilon-I, Sinergie, Lechner, Stiftelsen Norsk Institutt fof Luftforskning (NILU), Stichting Wageningen Research – Wageningen Environmental Research (WRWENR).

doi: 10.5281/zenodo.22231461

How to cite this report:
Mancosu, E., Plouvier, W., Sanchez, A., Milego, R., Dufek, J., Kleeschulte, S., Maucha, G., Kosztra, B., Milenov, P., & Petersen, J.-E. (2026). Summary report on potential methodological approach and resource needs for a regular EU ecosystem extent data product (ETC DI Report 2026/06). European Topic Centre on Data integration and digitalisation.

European Topic Centre on
Data integration and digitalisation (ETC DI)
https://www.eionet.europa.eu/etcs/etc-di

European Environment Agency
European Topic Centre
Data integration and digitalisation

Contents

List of figures

List of tables

Acknowledgements

The completion of this work was provided within the Negotiated procedure No. 3506/R0-COPERNCA/EEA.60044 between the European Environment Agency (EEA) and European Topic Centre on Data Integration and Digitalisation (ETC DI). The work was the result of Task 24: Support for ecosystem accounting & mapping, and it corresponds to the final deliverable D24.1. The report was redacted in close cooperation between the ETC partners and the EEA project team and benefited from external review by relevant experts in the field of remote sensing, ecosystem accounts and environmental policy implementation, namely Manuel Mayer, Bruno Smets, and Bruno Combal.

Description

Detailed and reliable data on ecosystem extent and distribution are critical to monitoring the achievement of legal obligations and policy goals in the European Union. This initiative aims to deliver a technical solution for a reliable, recurring data product for mapping ecosystem extent across Europe. Recent legislation, such as the amendment to the Environmental Accounts Regulation and the Nature Restoration Regulation (NRR), has further established the need for robust, spatially explicit ecosystem statistics and monitoring.

To meet these needs, the report evaluates different technical solutions: the traditional CORINE Land Cover (CLC) workflow and newer Copernicus Land Monitoring Service (CLMS) datasets. While CLC offers long-term consistency, CLMS provides higher spatial detail and more frequent updates.

The report concludes that an integrated approach is necessary, using the CLC workflow concept and logic as the workflow's foundation while incorporating the speed and thematic detail of modern CLMS data. The report also explores concrete examples of similar initiatives to support the deployment of a tailored approach that leverages Earth Observation (EO), artificial intelligence, and machine learning, while bringing human expertise and local knowledge for training and validation, in an incremental and co-creative manner.

Key recommendations for this strategy include adopting a tailored, streamlined approach to ecosystem characterisation and mapping, securing stable funding, ensuring a long-term cloud-based processing infrastructure, and developing a pan-European in-situ data strategy to support AI model training and independent validation.

Furthermore, the centralised production process is designed to reduce the administrative burden on individual countries by handling the most generic aspects of the technical work. Finally, this modernisation ensures the continuity of long-term data records, thereby preserving a vital geospatial data platform for land-use modelling and environmental indicators.

Executive summary

Recent policy developments have added urgency to developing a Europe-wide integrated ecosystem extent data product to fulfil both legal obligations and EU policy ambitions. This report brings together lessons from established mapping frameworks and the latest technological advances to propose a future-proof strategy for ecosystem mapping to support ecosystem accounting and assessment. It is primarily aimed at specialists and technical practitioners engaged in ecosystem accounting, spatial data production, remote sensing, and environmental policy implementation.

The report describes the outcomes of Copernicus User Task 24, "Support of Ecosystem Accounting & Mapping." The primary objective of that task was to propose a reliable, recurrent data product to support ecosystem accounting and mapping across Europe, ensuring full alignment with current EU legal and analytical requirements. The ecosystem data product developed and described herein focuses exclusively on mapping **ecosystem extent** across the European Union. It addresses the spatial distribution and areas of ecosystem types relevant to EU reporting and policy targets. This means it does not deal with spatial data for ecosystem condition parameters—such as soil organic carbon or urban green space indicators.

Recent EU legislation once more underlines the need for robust, consistent, and detailed spatial data on ecosystems. The Environmental Accounts Regulation and the Nature Restoration Regulation (NRR) lay the groundwork for recurring, pan-European reporting obligations, including requirements for spatially explicit ecosystem statistics and for monitoring restoration targets. Meeting these demands necessitates a product capable of mapping and tracking the spatial locations and areas of the diverse ecosystem types, in full coherence with the EU ecosystem typology developed for ecosystem accounting purposes.

The methodological review to design this data product considered two well-established technical solutions: the CORINE Land Cover (CLC)-based workflow and the integration of a diverse set of Copernicus Land Monitoring Service (CLMS) datasets. The CLC workflow, developed for consistency, comparability, and Member State engagement, has contributed decades of harmonised land cover and land use information. Yet, its traditional production cycles and coarse spatial resolution limit its responsiveness to recent policy needs. In contrast, CLMS datasets provide higher spatial detail and more frequent updates, but often lack coverage, thematic completeness, or harmonisation across the whole CLMS portfolio — posing challenges for achieving pan-European comparability and long-term continuity.

Through comparative assessment, clear insights emerge: no single approach alone is sufficient. The report recognises the robust foundation provided by the CLC workflow—especially its centralised coordination, endorsement, and coherence over time. Furthermore, the detailed class division of CLC (44 classes) closely aligns with the assessment needs of EU ecosystem accounts. Some CLC classes broadly indicate the presence of certain Annex I habitats from the EU Habitats Directive and can serve as a starting point for further mapping and field data collection. On the other hand, the newer CLMS data products demonstrate a faster, more efficient production process and can add useful thematic detail for certain ecosystem types. This report therefore recommends an integrated, tailor-made production workflow that builds on the CLC production approach as a foundational platform and adds useful experience and data from the modern CLMS production system.

This integrated workflow should accelerate update cycles to every three years, target finer spatial detail (aiming for a minimum mapping unit of 1 hectare for urban and 10 hectares for non-urban ecosystems, as a starting point), and increase available ecological detail. This means the ecological class division should support the second level of the EU ecosystem typology and, where relevant, its third level as well. Critically, the approach leverages recent advances in Earth Observation (EO), artificial intelligence, and machine learning and reasoning—especially for difficult-to-map ecosystem types such as grasslands and wetlands—but keeps human expertise and local knowledge at its core, ensuring data validation and credibility. The systematic inclusion of in-situ data remains indispensable for both training new models and rigorous, independent quality assessment.

Validation, transparency, and Member State participation are fundamental pillars in this vision. Rigorous multi-level validation—combining quantitative assessment against benchmarks and expert qualitative review—will be embedded in the production process. For its part, related to the land domain, the entire workflow will be documented and implemented within the remit of the CLMS, ensuring complementarity and integration with other CLMS data products. The intention of the centralised process is, to the extent possible, to relieve countries of performing the "mechanical part" of the technical work themselves, speed up data collection, and reduce administrative burdens.

The work under Task 24 has aimed to design a next-generation EU ecosystem mapping product that combines operational reliability, thematic depth, technological innovation, and EU – country level coordination. The Task 24 author team proposes the following ***key recommendations***:

1) Ongoing EU-level dialogues on strategic directions for the Copernicus Land Monitoring Service (CLMS) need to review options for securing stable funding for a CLMS product that effectively supports ecosystem mapping in support of the EU ecosystem accounting module, NRR implementation and EU-level assessments. It should foresee a wide-ranging operational functionality.
2) This review needs to consider funding for a centralised, iterative production model, ensuring the European Environment Agency (EEA) or another mandated entity has the mandate and resources to oversee the process. The design of the solution should consider, to the extent possible, the computing resources needed for its operationalisation, to allow the use of smart, modular, and scalable architecture.
3) Decisions need to be taken on the adoption of an upgraded and tailored CLC-based workflow as the reference method for ecosystem extent mapping at the EU scale, with binding minimum requirements for frequency, resolution, and thematic detail.
4) An essential complement would be the development and resourcing of a pan-European in-situ data strategy that builds on the experience of CORDA and the Copernicus In-situ Information System (CIS[2]). It should include standardised protocols, a common feature register and a protected database to support both training of AI models (used in-house or upcoming) and independent validation.
5) Mandate the alignment of future CLMS developments and national mapping efforts with these new standards, enabling seamless integration for both EU-wide and national reporting obligations. Create opportunities for countries to develop more detailed, locally specific ecosystem-related products and Copernicus downstream services.
6) Facilitate stakeholder engagement, particularly with national statistical offices and ministries, to ensure political buy-in, operational feasibility, and user acceptance.

Delivering a robust EU ecosystem data product will require strategic resource deployment and ongoing collaboration across organisational boundaries. Only when appropriate investment decisions are taken can ecosystem mapping data be delivered to meet EU policy needs, in support of the EU ecosystem accounting module, NRR implementation and EU-level ecosystem assessments. Furthermore, the traditional Corine land cover (CLC) dataset serves as a central geospatial data platform for developing a range of analytical products, including support for EU agri-environment indicators, land use modelling, and other geospatial analyses. Ensuring continuity of this data series in a modernised form would guarantee a continuation of the long-term CLC time series and provide a central geospatial data platform with diverse functionality for many other applications and EU indicators.

The ***report structure*** aims to lead readers stepwise from the underlying policy context and technical baseline to proposed solutions and operational guidance:

- **Chapter 1: Policy context and information needs**
  Presents the legislative background and evolving analytical requirements shaping the design of ecosystem extent datasets—highlighting the EU Ecosystem Accounting module, the Nature Restoration Regulation, and major reporting outputs such as the EU Ecosystem Assessment.

- **Chapter 2: Baseline assessment of mapping approaches**
  Critically examines two main technical options for EU-wide ecosystem extent mapping:
  - *Option 1:* CLC-based workflow incorporating incremental enhancements and AI-supported methods.
  - *Option 2:* Integration of established Copernicus Land Monitoring Service (CLMS) data products.

  Although each of the options contains common elements that could be shared, the entities of both options are mutually exclusive. Each approach is benchmarked across technical feasibility, thematic detail, spatial resolution, completeness, comparability, and operational continuity.

- **Chapter 3: Best practices and lessons from pan-European and national mapping projects**
  Reviews methodological innovations and implementation experiences from leading national and pan-European initiatives, including the latest advances in Earth Observation, object-based mapping, and AI integration.

- **Chapter 4: Recommended production strategy**
  Synthesises technical findings to propose a tailored production workflow. This chapter details:
  - Key enhancements over the established CLC method,
  - Workflow optimisation (process efficiency, automation, and collaboration with national experts),
  - Strategic use of AI and in-situ data,
  - Operational considerations for long-term continuity,
  - A fallback strategy based on direct CLMS integration as a contingency option.

- **Chapter 5: Conclusions and way forward**
  Summarises the overall assessment, highlighting lessons learned, core recommendations, and next steps for implementation and resourcing at the EU level.

Annexes
The annexes provide in-depth technical and methodological detail supporting the main recommendations:
- **Annex I:** Consistency and comparability of EU-wide land cover/land use (LCLU) products—methods and limitations.
- **Annex II:** AI enhancement of the CLC approach—opportunities and integration pathways.
- **Annex III:** Wetland nomenclature challenges—scope, harmonisation, and data gaps.
- **Annex IV:** Leveraging AI in European ecosystem mapping—potentials, requirements, and barriers.
- **Annex V:** Habitat mapping vs. modelling—concepts, methodological distinctions, and practical implications.
- **Annex VI:** The critical role of in-situ data in ecosystem extent mapping—foundations, protocols, and challenges.

1 Policy context and information needs

1.1 Overview

This chapter provides an overview of the information needs specified in EU legislative instruments that require data on the area and distribution of European ecosystems and habitat types. The most directly relevant EU directives and regulations include:

- The Ecosystem Accounting Module under the EU Environmental Accounts Regulation
- The Nature Restoration Regulation
- The LULUCF Regulation
- The Habitats Directive
- The Soil Monitoring Law

Only the first two legislative instruments are reviewed in detail in the following sections, as they most directly rely on comprehensive ecosystem extent data across the EU. However, this chapter also briefly addresses two important analytical outputs at the EU level that require spatial ecosystem information as a key input: the second EU ecosystem assessment ('MAES2IPBES') and the Global Biodiversity Framework indicator A2 ('Extent of natural ecosystems').

1.2 Ecosystem accounting: information needs

The amended Regulation (EU) No 691/2011 on environmental economic accounts requires EU Member States to report the area and changes in area for each ecosystem type within their national territory. These ecosystem extent accounts cover both terrestrial ecosystems (including freshwater) and marine ecosystems.

The EU ecosystem accounting module also requires reporting on a selected set of ecosystem condition indicators (not covered here) and a set of ecosystem service accounts. The latter depends on spatial information on ecosystem extent as a key analytical input. This module enables EU Member States to implement the global statistical standard on ecosystem accounting — the System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA).

Ecosystem types for extent reporting are defined according to the hierarchical EU ecosystem typology, with reporting at level 1 being mandatory and level 2 recommended. Reporting occurs every three years, starting with the reference year 2024. The reported data must represent a typical average for the reference year, expressed in thousand hectares and structured according to an agreed conversion matrix.

Ecosystem area is derived by summing the total extent of each ecosystem type. Ecosystem extent is defined as the size of an ecosystem asset projected orthogonally onto a two-dimensional plane. It is usually measured in spatial area using an appropriate coordinate reference system.

Accordingly, the resulting ecosystem data product can be considered a regular or irregular two-dimensional subdivision (tessellation) of the entire EU territory into contiguous, non-overlapping geometric units, each representing an individual ecosystem asset. These units, referred to as "basic spatial units" (BSUs) under SEEA conventions, may take the form of grid cells or other polygon geometries. According to Eurostat's technical specification for ecosystem extent accounting, the maximum size of a BSU should not exceed 1 hectare.

The SEEA EA requires that spatial data used to compile national ecosystem extent accounts must be *collectively exhaustive* — i.e. covering the entire national territory — and *mutually exclusive*, meaning each BSU must be assigned to only one ecosystem type at the specified typology level. This is known as the *MECE principle* (Mutually Exclusive and Collectively Exhaustive). The MECE principle and the logic behind ecosystem classification for accounting and assessment are embedded in the EU ecosystem typology, a nested hierarchical system that enables coherent class assignment across all three typology levels.

Although EU ecosystem condition accounts do not directly depend on spatial extent data, ecosystem service accounts do. Many of these service accounts require spatial data at level 2 of the EU ecosystem typology.

Other EU policies can also benefit from the ecosystem extent dataset, given their strong interconnections. In particular, the second EU ecosystem assessment and the EU Nature Restoration Regulation (NRR) are closely related. The NRR is examined in more detail in the following section.

1.3 EU Nature Restoration Regulation (NRR)

The Nature Restoration Regulation (NRR) is based on the conceptual and reporting framework of the Habitats Directive (HD) and sets targets for habitats that are not in a good state. Like the HD, the NRR requires regular reporting on the total area of all habitats and by habitat group (as defined in Annex I of the HD), separating out the share that is not in good condition. The scope of the NRR includes all habitats listed in Annex I, with a focus on those that are not common and widespread (i.e., covering more than 3% of a country's area). The minimum level of thematic granularity is the individual habitat group as defined in Annex I. Until 2030, priority will be given to habitats within Natura 2000 (N2000) sites.

While the HD does not explicitly define ecosystems, the NRR does. It defines an ecosystem as a dynamic complex of plant, animal, fungi, and microorganism communities and their non-living environment, interacting as a functional unit. This includes habitat types, habitats of species, and species populations. Since this definition is derived from the Convention on Biological Diversity, it largely aligns with the ecosystem definition used in the System of Environmental-Economic Accounting (SEEA). Consequently, an ecosystem extent data product would be able to provide information on the extent of habitats in the context of the NRR, as long as these habitats correspond to ecosystem types at Level 3 included in the product specifications. However, habitat condition itself will not be covered by the ecosystem data product.

The NRR includes specific provisions regarding the required information on the extent of certain terrestrial ecosystems, such as urban ecosystems, rivers, agricultural mosaics, mixed forests, and peatlands. The information needs for these NRR targets could be met by Level 2 of the ecosystem data product.
Additionally, the template for the Nature Restoration Plans includes a list of ecosystem types to be used as a reference when Member States define their restoration measures for terrestrial, coastal, and freshwater habitats of species. These must be organised by ecosystem type. The list includes:

a) Wetlands (coastal and inland)
b) Grasslands and other pastoral habitats
c) Rivers, lakes, alluvial, and riparian habitats
d) Forests
e) Steppes, heath, and scrub habitats
f) Rocky and dune habitats
g) Croplands
h) Urban habitats
i) Marine and coastal waters

In the NRR context, the list of ecosystem types provided in the uniform format for the nature restoration plans has been aligned to the extent possible with those from Level 1 of the EU ecosystem typology. This required an analysis and comparison of typologies (e.g., through crosswalk tables). Consequently, it appears that the NRR ecosystem list could be created by aggregating ecosystem types mapped at Level 2 of the EU ecosystem typology.

Annex IV of the NRR provides further details on how to report the share of agricultural land with high-diversity landscape features. It specifies that the following elements may be considered high-diversity landscape features, provided they meet certain management criteria:

- Land lying fallow, including temporarily
- Productive trees that are part of sustainable agroforestry systems
- Trees in extensive old orchards on permanent grassland
- Productive elements in hedges

The ecosystem data product could support the provision of spatial extent data for the above-listed features, if the following Level 2 and 3 ecosystem types are included:

- ET218 Fallow land
- ET24 Agroforestry, preferably subdivided into ET241 and ET242
- ET23 Permanent crops, including associated tree-dominant ecosystems at Level 3 (e.g., ET231 Olives, ET238 Hazelnuts, ET239 Chestnuts, etc.)

Since the NRR does not specify the minimum spatial detail and accuracy of the source data used to derive the extent of habitats and related ecosystems—as is the case with the HD—such information would be derived from relevant Copernicus Land Monitoring Service (CLMS) datasets (e.g., the Natura2000 monitoring product – N2K) or from dedicated habitat mapping projects within N2000 sites, such as EU Grassland Watch and the planned EU Wetlands Watch.

Table 1.1 provides a summary of the initial product specification for the habitat and ecosystem extent data product relevant to the NRR.

Table 1.1. Key requirements under the ecosystem accounting regulation (guidance note on ecosystem extent account) aligned with the requirements of the NRR

Key parameter	Value	Comment
Thematic detail (in relation to the ecosystem data product)	Level 2 (full coverage) Level 3 (selected habitat types only)	The main classification will be based on Level 2 of the EU ecosystem typology, enabling detailed ecosystem mapping and supporting the NRR. A subset of relevant Level 3 habitat types will be included for more detailed analysis, where applicable.
BSU type	Grid cell/polygon	The data may be represented as either regular grid cells or vector polygons, depending on the source data and the spatial representation needed
Spatial resolution	60x60 meters (at min)	High spatial resolution ensures a detailed and accurate ecosystem representation.
Spatial coverage (ecosystem accounting area)	EU 27	The data product covers the 27 EU Member States, ensuring comprehensive geographic coverage
MMU	- From EUGW (EU Grassland Watch) project 0.1 ha (post-2017) 0.9 ha (pre-2017) - From CLMS N2K: 0.5 ha	The smallest area that can be consistently mapped varies based on data source and time period. Recent data allow finer mapping (0.1 ha), while older datasets had higher thresholds. CLMS N2K datasets use a consistent 0.5 ha MMU.
Temporal reference	TBD	The exact reference year or period has not yet been defined. It will likely depend on data availability and alignment with reporting cycles
Update frequency	6 - 10 years	The data will be updated at longer intervals
Thematic accuracy	Overall accuracy of at least 85% (TBC)	A minimum accuracy threshold is proposed to ensure reliability, but final accuracy requirements are subject to validation and may be refined.

1.4 Second EU Ecosystem Assessment (MAES2/IPBES)

The first EU ecosystem assessment (Mapping of Ecosystems and their Services – MAES) was based on the MAES typology, which included nine main ecosystem types. This was later replaced by a typology of twelve ecosystem types to form Level 1 of the EU ecosystem typology for accounting, as referenced in Regulation (EU) No 691/2011 on environmental economic accounts. This regulation constitutes the only legal reference to an ecosystem typology within the EU. The original MAES typology is now considered a historical reference, though it may still appear in legislative proposals.

The second EU ecosystem assessment should, as far as possible, be consistent with the first assessment—conducted approximately five years ago—to enable evaluation of ecosystem changes over time.
The information needs of the second assessment regarding ecosystem extent are largely assumed to be addressed through ecosystem accounting. However, it is important to note that while ecosystem accounting emphasises temporal consistency of statistics and country-level reporting, ecosystem assessments require high-quality spatial representation of individual ecosystems and habitats across national and EU territories.

To ensure the ecosystem data product is optimally suited for the second ecosystem assessment and related policies, such as the Nature Restoration Regulation (NRR), the product specifications should include a spatial representation of the full range of Level 2 types within the EU ecosystem typology. **Table 1.2** provides an overview of the associated information needs.

Table 1.2. Key requirements under the ecosystem accounting regulation (guidance note on the ecosystem extent account) are aligned with the requirements of the 2nd EU ecosystem assessment

Key parameter	*Value*	*Additional details*
Thematic details	EU ecosystem typology level 1 and level 2	The product classifies ecosystems according to the official EU typology, using both broad (Level 1) and more detailed (Level 2) categories to ensure consistency and comparability.
BSU type	Grid cell	Each data point corresponds to a regular square grid cell, serving as the basic unit for spatial analysis and mapping.
Spatial resolution	100x100 meter	Ecosystem data are mapped at a spatial resolution of 1 hectare (10,000 m²), allowing for detailed spatial analysis across landscapes.
Spatial coverage	EU27 & EEA38	The dataset includes all 27 EU Member States, as well as other members and collaborative countries of the European Environmental Agency (EEA38), ensuring comprehensive geographic coverage
MMU	1 – 10 ha	The minimum size of ecosystem features that should be consistently mapped: 1 hectare in urban areas due to higher heterogeneity, and 10 hectares in non-urban areas
Temporal reference	Representative average for the reference year	The data reflect average ecosystem conditions during the reference year, providing a stable basis for temporal comparisons and trend analysis.
Frequency	3-year	The data product is updated on a triennial basis to capture ecosystem dynamics and inform regular policy reporting cycles.
Thematic accuracy	> 85%	The mapping aims to achieve a thematic classification accuracy of over 85%, which is considered the minimum acceptable threshold for policy-relevant spatial data.
Timeliness	Within 18 months after the reference year	Data must be available within 18 months of the end of the reference year, ensuring that countries can meet their legal reporting obligations and that Eurostat can conduct validations.

Completeness	In line with the MECE principle	The classification and mapping follow the "Mutually Exclusive, Collectively Exhaustive" (MECE) principle—ensuring no overlaps between classes and full spatial coverage

Global Biodiversity Indicator A.2 (‘Extent of natural ecosystems’)
Significant efforts have been made at the global level to ensure coherence between the Global Biodiversity Framework (GBF) and the UN System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA). As a result, the available guidance for Indicator A.2 of the Kunming-Montreal Global Biodiversity Framework (KM-GBF) recommends that the indicator on the *extent of natural ecosystems* be developed using national ecosystem extent accounts based on SEEA EA. This alignment makes the information needs of Indicator A.2 an additional important consideration in the development of a new ecosystem extent data product.

1.5 Other technical requirements

This section outlines additional parameters required to ensure the robustness of an Earth Observation (EO)-based product within a production framework. These parameters are also summarised in **Table 1.3**.

Thematic comparability
A robust EO-based product relies on thematic comparability of data products and underlying methods, including data collection approaches and integration techniques across different measurement instances. Maintaining consistency in input data sources, such as using the same type of satellite imagery, is also a significant advantage.

In the context of ecosystem accounting, this means applying uniform classification systems and standardised methods, particularly when assessing different geographic areas within the same observation period. Such methodological comparability directly translates into consistent outputs, enabling reliable, reproducible results with minimal methodological variation. When methods remain consistent across space, observed differences can be attributed to actual ecosystem changes rather than measurement inconsistencies. This enhances the credibility of the accounting system and supports meaningful comparisons.

Temporal consistency
In addition to thematic comparability, ecosystem accounting requires *temporal consistency* of datasets—allowing for comparisons over time within the same areas. While comparability supports consistent analysis across geographic regions, temporal consistency ensures that changes across time reflect real trends, not data noise or artefacts. Standardised accounting datasets—comprising cleaned, harmonised ecosystem extent data for at least two time points—are essential for accurately detecting temporal trends and changes.

Continuity concerns the temporal dimension of long-term time-series production. It involves developing methodologies that can withstand technological change. Long-term monitoring depends on the consistent availability of source data, which can be jeopardised by changes in satellite missions, sensor specifications, or processing workflows. A robust implementation ensures that data streams remain persistent while also allowing for adaptation to new sensors or improved algorithms—without compromising the consistency of the time series. This guarantees that observed changes reflect real environmental dynamics rather than methodological inconsistencies.

Transparency is another critical element. All methods, data sources, algorithms, and processing steps must be well documented and accessible, enabling others to fully understand and potentially reproduce the results. This documentation is a mandatory component of the final data product.

Finally, a critical requirement is the *validity* of the accounting data—its ability to accurately represent real-world conditions. This demands a dedicated validation workflow, separate from the production process. Validation compares the output classifications against reference data, ensuring the system measures what it is intended to measure.

Two main validation approaches are identified:

- **Quantitative assessment** provides a statistical comparison of the derived product against external benchmarks, helping evaluate scientific credibility and consistency across regions, often using official national reports.
- **Qualitative assessment** involves expert visual review ('look and feel' checks), which can identify systematic classification errors or confusion between ecosystem types by closely inspecting patterns and anomalies.

Both approaches contribute to a balanced validation framework, providing concrete insights into the reliability of the dataset and pinpointing areas where deviations occur.

Table 1.3. Additional requirements to contribute to the robustness of the ecosystem extent dataset

Parameter	***Value***
Consistency	Temporal consistency: Maintaining a robust methodological approach over time, allowing for accurate trend analysis and change detection.
Comparability	Possibility to compare the data with similar datasets for other geographic locations or countries, referring to seamless exchange and correct interpretation of ecosystem extent data across systems or organisational boundaries.
Transparency	Clear documentation of the dataset production, including the method description, is necessary to compile reporting metadata, as requested.
Continuity	Ensure long-term availability of the data source necessary to maintain the time-series production for monitoring and trend analysis
Validity	Quantitative and qualitative assessment

2 Assessment of ecosystem mapping approaches

Chapter 2 discusses two principal methodological approaches (CLC-based and CLMS integration) currently available for EU-wide ecosystem mapping. These approaches are assessed as reference options against both policy and technical requirements. The project's recommendations and proposals are developed in the subsequent chapters.

Given the limited capacity of current CLMS data to generate a comprehensive ecosystem extent account layer (see feasibility studies using CLC[1] and other CLMS[2] data), there is an urgent need to establish a viable solution to build a solid data foundation to exhaustively map the EU ecosystem typology at Levels 1 (EU L1) and 2 (EU L2), as required by the Environmental Accounts regulation.

In light of the findings from these feasibility studies, developing a new CLMS product appears to be the most practical solution. This production should align with the existing Copernicus production chain and build on established mechanisms, rather than rely on entirely new, untested techniques. Ongoing

[1] feasibility study based on CLC, is available to registered users.

[2] feasibility study based on CLMS is available to registered users.

evolution within the CLMS—through the introduction of new products and refinements to existing ones—should also be considered. At the same time, developments in the broader field of remote sensing may offer additional solutions for specific cases.

A key lesson learned from previous work is that, when refining a potential approach, input data should be selected from the existing CLMS portfolio—provided it is fit for purpose—rather than developing entirely new products. However, if existing products—even those with thematically relevant content—differ significantly from the intended objectives, it may be necessary to either substantially improve them or develop new datasets altogether.

Likewise, approaches that rely on input data from Member States (MS) must be approached with caution. The production workflow should avoid processes that require intensive harmonisation of highly heterogeneous national datasets, which may differ in terms of spatial coverage (data available from only a few countries) and thematic consistency (data classified using different systems).

Based on this premise, two options are explored and compared to determine the most suitable path forward:

- **Option 1** explores the efforts required to develop a dedicated workflow for producing an ecosystem extent layer directly from EO data. This would build on the existing CLC workflow while integrating new techniques to improve the process.
- **Option 2** evaluates the efforts needed to systematically gather and assess existing products already derived from EO data, in order to determine their suitability for the intended purpose.

This comparison aims to identify the most viable solution for designing an effective production workflow.

2.1 Main baseline options for ecosystem mapping

This section presents the two principal baseline approaches currently available for EU-wide ecosystem mapping. These operational pathways are evaluated as benchmark options against existing policy and technical requirements. However, they do not constitute the project's recommended solution.
The comparative assessment provided here serves as a reference point for further analysis. The project’s tailored recommendations and proposed solution will be developed and elaborated in Chapter 4.

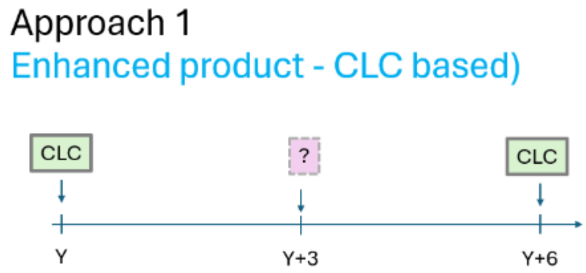
2.1.1 Option 1: CLC-based workflow with incremental improvements

Creating an EO-based workflow to derive an ecosystem map should not begin entirely from scratch. Over the past years, the European Environment Agency (EEA) has relied on the CORINE Land Cover (CLC) dataset for environmental accounting purposes, including ecosystem extent accounting.
Option 1, therefore, builds on the existing CLC production workflow, capitalising on its robust and harmonised land cover and land use data (**Figure 2.1**). This approach will require integrating innovative solutions that align with the EU ecosystem typology at Levels 1 and 2, as well as with the requirements outlined in the relevant Guidance Note.

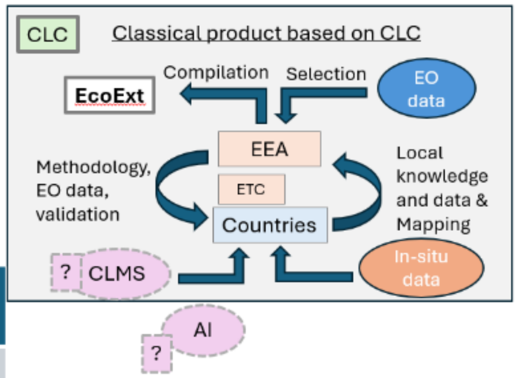
CLC is one of the most robust and established products for land cover and land use monitoring. Since the 1990s, the CORINE (Coordination of Information on the Environment) programme has implemented a standardised methodology for producing continent-wide land cover and biotope maps. Over the past three decades, CLC has been developed with the direct involvement of EEA member countries, resulting in a pan-European land cover and land use inventory consisting of 44 thematic classes, along with corresponding status and change maps.

The well-established methodology and the standard “change mapping first” approach used in CLC ensure time series consistency, spatial comparability, and the validity of mapped changes. Statistical consistency is further supported by the dedicated CLC accounting layers.

Figure 2.1. Schema of the starting point (CLC classical) for option/approach 1



Approaches	Fit for purpose	Spatio-Temporal coherence	Endorsement by countries	ROI
Assessment	Has limitations for the years between CLC-based instances	Poor since input products are not compatible	Depends on the acceptance of the approximation as reasonable compromise	No major investment. Relies on existing processes
Verdict				



While the existing workflow provides a solid foundation, it does not fully meet current ecosystem data needs—particularly in terms of delivering timely datasets with sufficient spatial accuracy. The primary limitation lies in the spatial resolution of the current approach. However, these shortcomings can be effectively addressed through a series of strategic enhancements.

The current CLC implementation already benefits from human intelligence, including visual evaluation, as well as valuable local knowledge and information contributed by national implementation teams. Building on this, the integration of recent and diverse satellite sensor data, the application of advanced Earth Observation techniques, the use of automated processes, and the deployment of AI capabilities can significantly improve the accuracy and timeliness of the resulting dataset—allowing the current constraints to be overcome.

2.1.2 Option 2: Integrated CLMS data workflow

Option 2 explores an approach based on the combination of existing EU-wide datasets (Figure 2.2). It does not involve extracting new or ancillary information from satellite or in-situ sources but instead relies exclusively on already harmonised data.

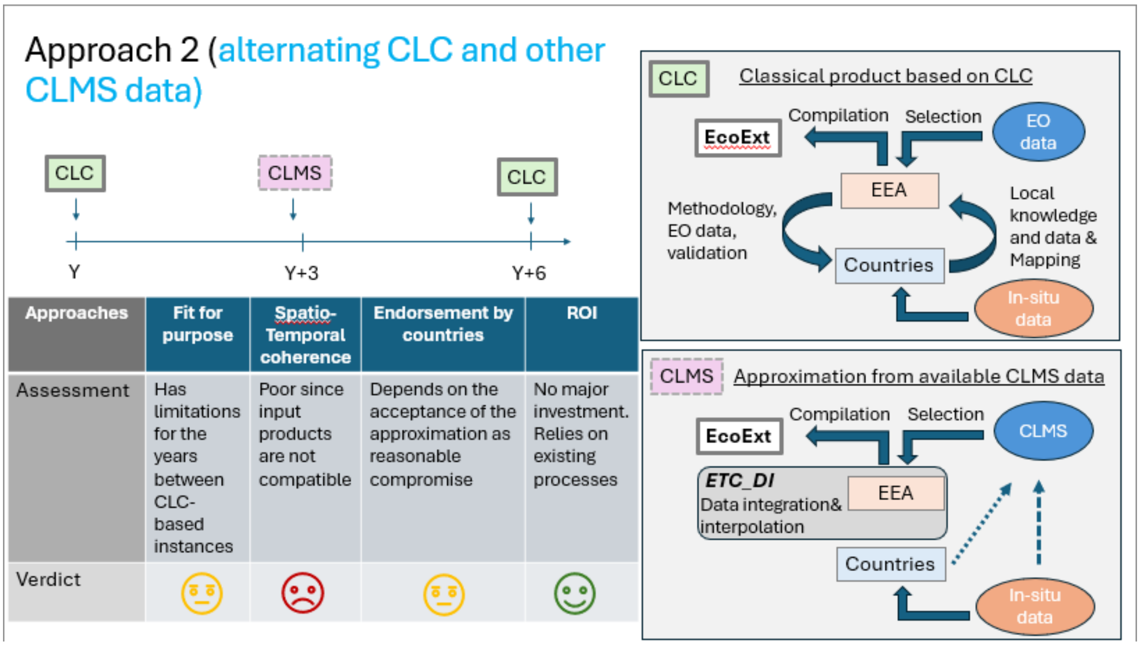
Drawing from previous experience, a clear example within this approach can be found in the existing CLMS portfolio—the most comprehensive and harmonised catalogue of land use and land cover data for monitoring Europe’s environment. To simplify the current analysis, other EU data sources are not considered at this stage. Nonetheless, this setup is sufficient to illustrate the multiple challenges inherent in a product-based integration approach.

The CLMS portfolio includes High-Resolution Layers in raster format at 10 m, 20 m, and 100 m resolutions. In parallel, the "Priority Area Monitoring" data is available in vector format, providing a high level of spatial detail with geometries[3] typically having a Minimum Mapping Unit (MMU) of less than 1 hectare.

The **Corine Land Cover Accounting Layers (CLCACC)** [4] offer a coherent and reliable foundation for this approach. These raster time series at 100 m resolution can be enhanced by integrating other datasets, which are often available at finer spatial resolutions.

However, the main challenge lies in ensuring consistent integration of these diverse products, which often differ in nomenclature, production timelines, reference years, and spatial resolution.

Figure 2.2. Product-based integration – schema of option/approach 2



2.2 Comparison of options

This chapter provides a critical review of operational solutions and workflows currently available or recently piloted for EU-scale ecosystem extent mapping. While these approaches are compared against the project's requirements, they do not represent the recommendations of this report.
The section explores in detail the key parameters underlying these requirements to identify the strengths and limitations of the two main approaches considered for creating an ecosystem data product. The description of these approaches reflects the current conceptual state of the art, drawing on evidence and insights from prior experience.

This analysis provides a foundation for understanding what is needed to streamline future ecosystem data-monitoring workflows. By highlighting the advantages and limitations of each approach, it clarifies the areas that require further development. It is important to note that this evaluation is based on the information available at this stage and may not yet be exhaustive.

[3] The classified spatial objects according to the product classification/nomenclature
[4] https://www.eea.europa.eu/en/datahub/datahubitem-view/a55d9224-a326-4cb1-9b9c-3a324520341a

2.2.1 Comparative criteria: thematic content and nomenclature

The ecosystem data foundation must align with the nomenclature defined in the technical note on the EU ecosystem typology (Eurostat, December 2024). This typology has been endorsed as a common classification framework to harmonise reporting on ecosystem extent and condition across EU Member States (MS).

The classification is structured into three hierarchical levels, with **Level 1 (EU L1)** and **Level 2 (EU L2)** serving as the primary targets for the ecosystem extent data product. **Level 3 (EU L3)** offers an additional, more detailed set of ecosystem types to support ecosystem service, condition modelling, and national uses.

- **Level 1 (EU L1)** consists of 12 sub-types and is derived from the MAES (Mapping and Assessment of Ecosystem Services) ecosystem typology.
- **Level 2 (EU L2)** includes 47 sub-types and incorporates more detailed land use information (e.g., CORINE) and habitat classifications (e.g., EUNIS Level 2).
- **Level 3 (EU L3)** includes 145 sub-types and integrates further detail on crop types and plant species. It aims to represent EUNIS habitat types where feasible.

The EU ecosystem accounting regulation mandates L1 and recommends L2. These two upper levels define the scope of the ecosystem data foundation.

This scope may be expanded in the future by incorporating selected **L3 types** that provide high added value. For instance, **ET 2.3.11** (Semi-natural elements associated with agricultural land use in permanent crops), under **ET 2.3** (Permanent Crops), is relevant in the context of understanding the ecosystem condition of this L2 ecosystem type. Furthermore, incorporating L3 types may enhance consistency with habitat reporting obligations under the Nature Restoration Regulation (NRR).

General considerations

Despite the growing availability of satellite data, in situ observations (e.g., Copernicus In-Situ Component), and advances in data processing and interpretation, the scientific understanding and detailed mapping of ecosystems to meet environmental policy monitoring requirements remain a significant challenge in certain areas.

In this specific context, producing a data product capable of capturing the extent of the 47 classes defined in the EU Ecosystem Typology at Level 2 (ET L2) requires more than the photointerpretation of optical imagery alone. Additional data on land use, management practices, and key biophysical elements—such as soil and climate—are essential. While some of this information can be partially derived from optical sensors, the use of other technologies, such as radar or in-situ measurements, is strongly recommended. In-situ data, which reflect real-world conditions on the ground, are crucial for understanding, quantifying, and analysing land states and processes. They enable more accurate observation and interpretation of ecosystems, complementing satellite-derived data.

Table 2.1 provides an initial overview of the key challenges in accurately mapping specific ecosystem classes, grouped by EU Level 1. This overview is not exhaustive but aims to offer a general understanding of the challenges within the ET nomenclature. The 'Accuracy level' column presents a feasibility-based estimate, informed by prior experience in constructing each ecosystem type using various input datasets.

Table 2.1. Summary table of critical Land Cover / Land Use elements to map ET classes

ET	Mapping challenges of Ecosystem classes at Level 1 and Level 2	*Mapping feasibility & Accuracy level*
ET1 Settlement	Recent advancements in the CLMS portfolio and other initiatives have demonstrated significant improvements in mapping artificial areas. However, the ET1 requires high-detail mapping to accurately differentiate diverse types of buildings and infrastructure, along with the associated open spaces, as exemplified in CLC or Urban Atlas. Specifically, the main challenges are to distinguish between infrastructures (ET 1.3) and urban/residential (ET 1.1 and 1.2), and specific urban elements, such as dump areas (**ET 1.3**), green urban areas and leisure areas (**ET 1.4**), and cemeteries (**ET 1.5**).	*Moderate-high*
ET2 Cropland	Good achievements in classifying main crop types, such as those characterised by clear vegetation structure, as permanent crops (**ET 2.3**) or rice fields (**ET 2.2**), as well as those with predictable seasonal harvest patterns (**ET 2.1**). The complexity is within mixed agriculture patterns (**ET 2.5**), which might include woody permanent crops and different crop types, as well as within agroforestry areas (**ET 2.4**)	*High*
ET3 Grassland	Grassland ecosystems are characterised by heterogeneous vegetation types that vary across different geographical locations, making accurate mapping a challenge. Within the ET, in alignment with CLC, a division between natural/semi-natural grassland (**ET 3.2**) and modified grassland (Sown pastures and other grass) (**ET 3.1**) is requested. To successfully distinguish these two main grassland types, the mapping should combine multi-temporal optical and radar imagery with ancillary data on topography, climate, and land management practices.	*Moderate*
ET4 Forest	Tree cover classification, focusing on broadleaved (**ET 4.1**) and coniferous (**ET 4.2**) types, and even evergreen broadleaved (**ET 4.3**) is relatively straightforward compared to more detailed typologies, such as other forest types. A general limitation in forest ecosystem mapping is the inclusion of forestry elements as transitional forest (**ET 4.5**) and temporarily unstocked forest, such as clear-cut areas where the trees do not yet meet the minimum height, or the crown coverage as typically included in forest definition. Analysis of image time series can help identify unstocked forests. A particular challenge is the identification of plantations (**ET 4.6**)	*High*
ET5 Heathland	Complex class, usually represented by mixed vegetation structure, may cause potential confusion with other vegetated lands; sparsely vegetated and grassland are common. Heathlands typology presents challenges in mapping specific categories like tundra (**ET 5.1**), due to its unique vegetation patterns and climatic conditions. The integration of multiple factors, such as in-situ data and biophysical information, may be required to distinguish heathland/ tundra from grassland and sparsely vegetated ecosystems.	*Moderate*
ET6 Sparsely Vegetated	Sparsely vegetated ecosystems class presents mapping challenges from their low degree of vegetation coverage and the presence of non-vegetated ecosystems such as bare lands, rocks (**ET 6.1**) and semi-desert and sparsely vegetated areas (**ET 6.2**). To achieve accurate results, multi-temporal imagery from both optical and radar sensors should be combined with in-situ information, such as geological or topological maps, among others (providing they are available and accessible).	*Moderate*
ET7 inland wetlands	The complexity of detecting wetlands is related to the seasonality of some wetland types, which alternate between water, bare land and vegetated land covers. Another challenge is detecting specific wetlands associated with specific soil characteristics, such as peatlands, which requires ancillary information of soil type and in situ information to accurately map mires, bogs, and fens (**ET 7.2**).	*Moderate*

ET8 Rivers	Mapping the water bodies, such as rivers characterised by flowing water, is straightforward, especially when DEM data is used as additional support. The main challenges exist when distinguishing between natural rivers (**ET 8.1**) and artificial channels (**ET 8.2**)	*Moderate-high*
ET9 Lakes	Mapping water bodies, such as lakes characterised by still waters, is straightforward. However, complexity arises when differentiating specific lake categories, like artificial reservoir (**ET 9.2**) and geothermal pool (**ET 9.3**).	*High*
ET10 Marine inlets	Coastal lagoons (**ET 10.1**), Estuaries (**ET 10.2**) and intertidal flats (**ET 10.3**) are characterised by the interactions between water and land, driven mainly by tidal fluctuations. Their boundaries can be influenced by sediment dynamics, and the presence or absence of vegetation helps distinguish them from other coastal ecosystem types. In-situ data are needed to identify the boundary of tidal influence or seawater.	*Moderate-high*
ET11 Coastal	The complexity of the class consists in mapping coastal wetlands as coastal saltmarshes and salines (**ET 11.4**), along with sparsely vegetated ecosystems in close proximity to marine water, such as coastal dunes, beaches, sandy and muddy shores (**ET 11.2**) and rocky shores (**ET 11.3**) situated above the high-water mark. The greatest challenge of this class is mapping man-made dykes and dams, including wave breakers extending into the sea, constructed primarily to protect land from the sea (**ET 11.1**).	*Moderate-high*
ET12 Marine	Marine realm typology mapping presents significant challenges, as demonstrated by various initiatives; it requires a combination of different EO sensors and in-situ information to accurately map marine ecosystem types, as demonstrated by the seabed habitat map produced by EMODnet (https://emodnet.ec.europa.eu/en/seabed-habitats).	*N/A*

Building on the initial findings presented in **Table 2.2**, **Table 2.3** provides a more detailed analysis of the feasibility of mapping specific ecosystem types, categorised into three levels: **high**, **moderate**, and **low**.

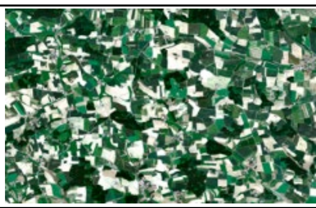
- **High feasibility** applies to ecosystem types that can be reliably detected through manual (computer-assisted) photointerpretation of optical and radar imagery, with little or no need for ancillary data. These ecosystem types exhibit stable, clearly observable biophysical and morphological characteristics across time and space, enabling unambiguous identification and mapping. Their features can be captured effectively from a single satellite image—provided it has adequate spatial, spectral, and radiometric resolution—or from a targeted set of images acquired at specific times of the year. In some cases, a denser time series may be required to capture seasonal dynamics. With a sufficient supply of training data (e.g., photo-interpreted samples and in-situ data), the process of capturing these ecosystem types via EO can potentially be automated.
- **Moderate feasibility** refers to ecosystem types that also display observable biophysical and morphological characteristics but are not distinct enough for unambiguous detection using EO data alone. Mapping these ecosystems typically requires the integration of ancillary datasets, such as land use information or soil maps, to improve accuracy and confidence. In these cases, the EO signal may be too general to distinguish among similar ecosystem types, or it may vary significantly within the same ecosystem class due to local land management practices. The availability of relevant in-situ data can help to refine and enhance the EO-based mapping process.
- **Low feasibility** characterises ecosystem types that lack distinct biophysical or morphological features that can be consistently detected using EO data. These ecosystems cannot be reliably mapped solely through remote sensing. Instead, mapping relies heavily on the direct ingestion and integration of in-situ data, either alone or in combination with EO imagery. Manual photointerpretation may also be required. The strong dependence on in-situ information for capturing specific biophysical characteristics makes these types particularly difficult to identify using EO-based methods.

The photointerpretation approach used in **CORINE Land Cover (CLC)** relies on the visual interpretation of physical features—representing real-world phenomena—as observed from one or more satellite images. Two essential conditions must be met for this method to succeed:

1. The ecosystem must possess **biophysical and morphological characteristics** that clearly and consistently define it and that are observable from space.
2. The satellite data used must have the **technical specifications** (spatial, spectral, and radiometric resolution) necessary to reliably detect those characteristics.

Table 2.2. Feasibility assessment by ET classes considering the typology of input data and sources

	EU ecosystem typology Level 2 - Feasibility		
	High	**Moderate**	**Low**
	Detectable through the information in the satellite image. Some ecosystem types may require indices derived from EO data and/or time-series analysis.	Ecosystem types that cannot be differentiated solely based on their biophysical and morphological characteristics detectable on satellite images, but rather require ancillary geoinformation.	Unfeasible to be detected by remote sensing. Specific in-situ data sources depicting the extent of the ecosystem are required.
ET1 Settlement			
	1.1 Continuous settlement area & 1.2 Discontinuous settlement area: Urban areas can be easily mapped by their spectral signal and morphological characteristics.	**1.4 Urban greenspace**: Areas with vegetation within or partly enclosed by urban fabric exhibit a complex spectral signal due to the combination of artificial structure and natural vegetation patterns. Still, the green areas could be captured reasonably well with high- or very-high-resolution satellite data. However, the correct assignment of a green patch or water body as part of urban greenspace is difficult without additional information on the location and spatial context. This makes the mapping challenging without in-situ data on green/blue areas associated with urban areas. The position of such areas in local context (green patch within the dense urban fabric) could be used for their identification as part of the "urban green".	**1.3 Infrastructure and industrial areas**: Remote sensing allows for the detection of urban structure, but its use cannot be derived unless it is clearly evident from the morphological characteristics. Ancillary geoinformation on infrastructure is required, and it is relatively scarce. Special difficulty in distinguishing mines/dumps from ET6, sparsely vegetated ecosystems.

			
			1.5 Other artificial areas: This class includes, e.g. permanent greenhouses, cemeteries, even if predominantly green. Similar to class 1.4, the mixed composition of artificial and vegetative elements requires additional in-situ data on small-scale urban features (that most often do not exist).
ET2 Cropland			
	2.1 Annual cropland: This class can be distinguished by its spectral characteristics and the analysis of time series (seasonality).	**2.3 Permanent crops:** The spectral characteristics of this class can be very similar to other vegetation classes. Ancillary information on specific vegetation structure can help to detect this class.	**2.4 Agro-forestry areas:** This ecosystem type can be characterised by the presence of different land cover elements, such as crops and trees. Due to the heterogeneous composition of this class, detection using Remote sensing alone is not feasible. In-situ data on agro-forestry management practices is essential for accurate mapping.
			
	2.2 Rice fields: This class can be distinguished by its spectral characteristics and the analysis of time series (to differentiate from wetlands when flooded).	**2.5 Mixed farmland:** This class includes a combination of other classes (none of them dominant), and detecting it depends on the study scale.	**2.6 Other farmland:** This class combines different types of farmlands. Requires in-situ data on farmland management practices (that most often do not exist).
ET3 Grassland			
		3.2 Natural and semi-natural grasslands: This class can be detected by its spectral signal and spatial distribution, although variation occurs in different geographical locations.	

			
		3.1 Sown pastures and other grass (modified grasslands): Remote sensing allows to detected grasslands, but their use cannot be derived (unless pairing their temporal behaviour with local information on farming practices). Requires ancillary geoinformation on modified grasslands.	
ET4 Forest			
	4.1 Broadleaved deciduous forest: This class can be distinguished by its spectral characteristics.	**4.4 Mixed forest:** This class includes a combination of other classes' areas (none of them dominant), and detecting it depends on the study scale.	**4.5 Transitional forest and woodland shrub:** This class combines different types of forestry elements. Requires in-situ data on land management practices and time series, especially for detecting temporarily unstocked forests.
			
	4.2 Coniferous forest: This class can be distinguished by its spectral characteristics.		**4.6 Plantations:** Remote sensing allows to detect forested areas, but their type/origin cannot be derived. Requires in-situ data on forest management practices (that most often do not exist).
			
	4.3 Broadleaved evergreen forest: This class can be distinguished by its spectral characteristics.		

ET5 Heathland			
		5.1: Tundra: Unique vegetation pattern and climatic conditions. This class can be distinguished by its spectral characteristics, but ancillary biophysical information & time series are required, and ancillary data are preferable.	**5.2 Scrub and heathland:** The spectral characteristics of this class can be very similar to other vegetation classes. Requires image interpretation and preferably ancillary data.
			
		5.3 Sclerophyllous vegetation: Unique vegetation pattern and climatic conditions. Ancillary biophysical information is required, and preferably, ancillary data.	
ET6 Sparsely vegetated			
	6.3 Ice sheets, glaciers and perennial snowfields: This class can be distinguished by its spectral characteristics and the analysis of time series.		
			
	6.1 Bare rocks: This class can be distinguished by its spectral characteristics, but confusion can occur with other sparsely vegetated classes. Requires image interpretation and the analysis of time series.	**6.2 Semi-desert, desert and other sparsely vegetated areas:** The spectral characteristics of this class can be very similar to other sparsely vegetated classes. Requires image interpretation and preferably ancillary data.	

ET7 Inland wetlands			
		7.1 Inland marshes and other wetlands on mineral soil: This class can be distinguished by its spectral characteristics, but ancillary soil type data are required, and in-situ data are preferable.	**7.2 Mires, bogs and fens:** The spectral characteristics of this class can be very similar to other wetland classes. Requires image interpretation and preferably ancillary soil type data.
ET8 Rivers			
	8.1 Rivers and streams: This class can be distinguished by its spectral characteristics. Additional information from DEM and the analysis of time series to consider seasonal and temporal changes in water level in rivers that present great seasonal changes.	**8.2 Canals, ditches and drains:** The spectral signal of this class is like that of other classes. Requires in-situ data on human interference.	
ET9 Lakes			
	9.1 Lakes and ponds: This class can be distinguished by its spectral characteristics. In some cases, ancillary data will be needed.	**9.2 Artificial reservoirs:** The spectral signal of this class is like that of other classes. Requires in-situ data on human interference.	
			
		9.3 Geothermal pools and wetlands (Iceland): Unique climatic conditions. The spectral signal cannot be differentiated from other classes; thus, ancillary data are required.	

ET10 Marine inlets			
	10.1 Coastal lagoons: This class can be distinguished by its spectral characteristics. Confusion possible with ET 12 Marine.	**10.2 Estuaries and bays:** This class can be distinguished by its spectral characteristics, although difficult to define limits with rivers. It can be confused with ET 12 Marine and ET8 Rivers.	
			
	10.3 Intertidal flats: The spectral signal of this class is like other classes. Need for time-series information to account for tidal fluctuations.		
ET11 Coastal			
	11.2 Coastal dunes, beaches and sandy and muddy shores: This class can be detected by their spectral signal, although variation occurs in different geographical locations. Confusion possible with 11.4 Coastal salt marches and salines & 10.3 Intertidal flats	**11.4 Coastal salt marshes and salines:** The spectral signal of this class is similar to that of other classes. Salines are not always square features and are hard to identify. Ancillary soil type data, salinity information (Salinity Index), are required, and in-situ data is preferable.	**11.1 Artificial shorelines:** The spectral signal of this class is like that of other classes. Requires in-situ data on human interference near shorelines. Limitations are also due to the limited size of artificial shoreline elements, such as Man-made dykes, dams, or wave breakers, which makes their detection challenging.
			
	11.3 Rocky shores: This class can be detected by their spectral signal, although variation occurs in different geographical locations.		

2.2.2 Comparative criteria: nomenclature gaps and alignment

Upon examining the two proposed approaches—one focused on enhancing the CLC workflow and the other on integrating existing data across the EU—it becomes clear that current information is insufficient to fully align with the EU ecosystem typology at Levels 1 and 2. This sub-chapter aims to identify existing gaps and key constraints that may hinder the creation of a comprehensive, wall-to-wall ecosystem dataset.

In the first approach, the CLC provides a reasonable estimate for several classes required at the Ecosystem Typology Level 1. However, it also presents notable limitations. One of the main issues stems from the relatively large Minimum Mapping Unit (MMU) used in the CLC process. As a result, several required ecosystem classes are not captured as distinct units but are instead aggregated within broader CLC categories. This makes them difficult to isolate and accurately assign to the appropriate ecosystem typology class.

In the second approach, an additional distinction was made between including and excluding CLC data alongside other CLMS data during the integration process. This distinction aims to illustrate the differences in results that may arise and to emphasise the capacity of newer CLMS products to provide greater thematic detail compared to the CLC, even though several classes are not fully covered due to the limited geographical scope of some products. Nevertheless, this approach highlights successful cases of mapping classes absent in the CLC, such as broadleaved evergreen forests, while other features—such as 'geothermal pools' or 'other farmland'—remain unmapped.

Table highlights the key elements that cannot currently be distinguished or correctly mapped in line with the EU typology, based on the available CLMS portfolio.

- **Classes highlighted in yellow** (e.g. greenhouses, in the CLMS without the CLC column) are present in certain CLMS products but lack complete coverage across the European territory.
- **Classes highlighted in red** (e.g. greenhouses in the CLC-based column or artificial shorelines) are either entirely missing or incorrectly classified within the EU ecosystem typology.

Specifically, some of these classes should correspond to ET1 and ET11 in the EU typology, but under the current classification system, they are misassigned—for example, greenhouses are mapped under ET2, and artificial shorelines under ET1, which is incorrect.

Table 2.3. Key classes (or elements) that are missing for ET mapping across the two approaches

	Status of the classes or element in the different approaches: partially available yellow*; missing: red		
	Current CLC Workflow with Incremental Improvements	Integrated CLMS Data Workflow	
		CLMS without CLC	CLMS with CLC
ET1	1.5 Other Artificial areas (Greenhouses, cemeteries)	1.5 Other Artificial areas (Greenhouses*, cemeteries[5])	1.5 Other Artificial areas (Greenhouses*, cemeteries)
ET2	2.6 other farmland	2.4 Agro-forestry areas*; 2.5 Mixed agricultural patterns*, 2.6 Other farmland	2.6 Other farmland*
ET3		3.1 Modified grassland*, 3.2 Natural grassland*	

[5] Available in OpenStreetMap (OSM)

ET4	4.3 Broadleaved evergreen; 4.6 plantations	4.5 Transitional forest*; 4.6 plantation *	4.6 plantation*
ET5	5.1 Tundra	5.2 heathland*, 5.3 Sclerophyllous vegetation*	
ET6		6.1 bare rocks*, 6.2 semi-desert and other sparsely vegetated*	
ET7		7.1 Inland marshes*, 7.2 mires, bogs and fens*	
ET8	8.2 Artificial watercourses	8.1 Rivers*, 8.2 Artificial watercourses*	8.2 Artificial watercourses*
ET9	9.2 Artificial reservoirs, 9.3 Geothermal pools	9.1 Lakes*, 9.2 Artificial reservoirs*, 9.3 Geothermal pools	9.2 Artificial reservoirs*, 9.3 Geothermal pools
ET10		10.1 Coastal lagoons*, 10.2 Estuaries and bays*, 10.3 Intertidal flats*,	
ET11	11.1 Artificial shorelines	11.1 Artificial shorelines, 11.2 Coastal dunes*, 11.3 Rocky shores*, 11.4 Coastal saltmarshes and salines*	11.1 Artificial shorelines

It is important to note that these comparisons do not account for the required Minimum Mapping Unit (MMU), nor the frequency or timeliness of production. Instead, the aim is to demonstrate the feasibility of mapping thematic classes in CLMS to the required nomenclature.

2.2.3 *Comparative criteria: frequency, MMU, and spatial resolution*

In general, the CLMS data products strive to align well with target requirements for frequency, minimum mapping unit (MMU), and spatial resolution. In contrast, the current CLC workflow—despite some improvements—currently does not meet these specified requirements. However, as shown in **Table 2.4** , even the CLMS-only approach only partially fulfils the criteria. This partial compliance is due to the irregular temporal frequency and variable MMU of the CLMS input layers, as well as inherited limitations from using CLC data, which typically features a lower frequency and a larger MMU than required.

Based on experience with CLC, additional considerations must be made when different MMUs are used within a single data product. Employing varying MMUs in 'CLC status' and 'change layers' introduces several technical and statistical challenges. Enhancing CLC suitability by applying a 5-hectare MMU would demand significant effort but could greatly improve the quality and content of the delivered information and resolve many longstanding logical issues between status and change layers. High-quality accounting layers could then be efficiently generated by combining the enhanced status with existing change data.

The choice of a 100-meter raster resolution for CLC reflects a traditional compromise that, in some cases—such as representing linear elements like road networks—may lead to information loss. Technically, it is possible to rasterise the original CLC vector data at a finer resolution and produce an accounting version as needed.

Table 2.4. Frequency, MMU and spatial resolution comparison

Legend of colours	
(green)	Meet the requirements
(yellow)	Partially meets
(pink)	Does not meet

	Current CLC Workflow with Incremental Improvements	Integrated CLMS Data Workflow
Frequency	The frequency process of CLC follows a six-year cycle, which does not align with the specific requirements	Several data products from the CLMS portfolio support a frequency aligned with the requirements. However, the frequency of the different CLMS products is not harmonised, and crucial inputs like the CLC are updated less frequently than the target requirement.
MMU	Within CLC, the land cover/land use information is mapped with an MMU of 25 ha (5 ha for the change layers).	Most of the data products from the CLMS portfolios are available within the MMU, as indicated by the target requirement.
Spatial resolution	100m	Less than 100 m

2.2.4 Comparative criteria: temporal consistency and thematic comparability

Ensuring consistency over time is one of the greatest challenges in EO-based land cover and land use monitoring products. Temporal consistency is essential for accurate change detection and trend analysis, as the accounting module relies on comparable input data to compile meaningful accounting tables. Although recent initiatives have improved spatial resolution and update frequency, achieving consistency across different releases is often overlooked.

The current CLC workflow, incorporating incremental improvements, addresses this issue by introducing additional “accounting layers” alongside the status and change layers produced every six years. These additional layers are modified status layers specifically designed to guarantee consistent statistical analysis within the EEA’s land cover change accounting system. However, due to the technical characteristics of CLC and CLCC data, the evolution of the CLC production methodology, and varying quality of input data, time-series statistics derived directly from historical CLC data exhibit several inconsistencies. To establish a robust statistical basis for CLC-based time-series analysis, a harmonisation methodology has been developed.

Consistency and comparability in CLC are further supported by centralised coordination under the EEA, which includes providing standardised satellite image inputs, detailed technical guidelines, dedicated software tools, training for national teams, and thematic verification. This comprehensive approach ensures a harmonised output from Member State mapping activities.

In contrast, other CLMS products currently lack systematic mechanisms to ensure full temporal harmonisation. For example, apart from CLC, other Copernicus products do not provide a systematic production of change layers, which are essential for helping users distinguish genuine land cover changes from artefacts caused by methodological updates.

A current limitation of CLMS is the heterogeneous geographic coverage of some datasets, which results in certain areas having greater thematic detail than others and leads to potential imbalances in spatial consistency.

Moreover, when integrating multiple datasets, it is important to assess potential inconsistencies in comparability. This requires additional analysis to determine how consistently different products map the same features before accurate information can be extracted. Priority rules also need to be established to resolve conflicts when different input data sources provide contradictory information for the same location. Annex I provides a summarised overview focusing on the limitations related to thematic comparability and temporal consistency in LC/LU information within CLMS data.

As shown in **Table 2.5**, an approach based on a centralised workflow is more likely to meet the requirements than an approach relying on multiple inputs, which often lack harmonisation for temporal and regional comparability, and are therefore less effective. Option 2 would require additional efforts to implement effective mechanisms that ensure greater comparability and consistency.

Table 2.5. Consistency and comparability comparison

Legend of colours	
(green)	Meet the requirements
(yellow)	Partially meets
(red)	Does not meet

	Current CLC Workflow with Incremental Improvements	Integrated CLMS Data Workflow
Consistency and comparability	An approach based on a centralised workflow and CLC's 'Change mapping first approach' methodology can effectively support the estimation of real changes with relatively low effort. As learned from the CLC production, temporal consistency is effectively addressed with the accounting layers, which ensure statistically comparable results. These layers mitigate inconsistencies that may arise from changes in the methodologies and source data.	Production set-ups that integrate multiple inputs from likely independent initiatives typically yield unstable results. Such an approach would require corrective 'accounting' methodologies across different production cycles, which becomes increasingly challenging.

2.2.5 Comparative criteria: completeness

The current CLC workflow provides a strong foundation for meeting completeness criteria. As demonstrated by its widespread application in previous work, the CLC nomenclature offers information across three hierarchical levels and adheres to the 'MECE principle' (Mutually Exclusive and Collectively Exhaustive).

In the context of the integrated approach and current CLMS data, it has been shown that integrating multiple datasets with similar thematic coverage presents challenges due to the heterogeneity of information. An additional limitation is that, when combining multiple sources for the same area, there is a risk of the opposite problem—some areas may end up lacking information altogether. As indicated in

Table 2.6, the first approach is better positioned to satisfy the requirement, whereas the second approach will require further measures to ensure comprehensive coverage across the entire spatial extent.

Table 2.6. Completeness comparison

Legend of colours	
(green)	Meet the requirements
(yellow)	Partially meets
(pink)	Does not meet

	Current CLC Workflow with Incremental Improvements	Integrated CLMS Data Workflow
Completeness	A centralised approach, such as the CLC, ensures harmonised understanding and application of nomenclature, and workflow CLC nomenclature allows systematic allocation of unique information to each class.	An approach that integrates multiple available inputs may face greater challenges in achieving completeness. It would require control phases, guidance for the member states and corrective adjustments due to the heterogeneity of information, which is created to serve multiple purposes.

2.2.6 Comparative criteria: continuity and operational dependency

The production of EO-based data must consider dependency factors during the planning phase, as these can significantly influence the success of future releases. As highlighted previously, the data foundation underpinning the ecosystem accounting regulation requires that the product be purpose-driven and designed for repeated production over multiple cycles—ideally within a three-year interval—rather than as a one-off output.

Therefore, it is essential to identify potential dependency factors during planning, as they play a crucial role in ensuring the smooth realisation of the final product. Ensuring the continuity of an ecosystem data product depends on consistent data acquisition, stable methodologies, coherent classification systems, and the involvement of qualified experts to guarantee the correct application of procedures. All procedures should be thoroughly documented to maintain transparency and reproducibility throughout the product lifecycle.

The creation process involves various data sources, and the more diverse the inputs and linkages, the greater the risks associated with production. The different inputs considered in the Integrated approach—though ideally all from the CLMS portfolio—demonstrate a degree of heterogeneity in methodology and objectives. Building ecosystem data products based on these inputs means relying on multiple data sources, methods, and experts, which may potentially diverge from the initial setup. This approach requires well-defined production criteria to ensure continued alignment with the original objectives.

An improved approach that utilises the CLC workflow—while retaining its fundamental principle of centralised technical coordination by the EEA, with production by Member States and continued support for funding and governance—proves advantageous for maintaining continuity in an ecosystem product. The current setup has demonstrated long-term reliability, even as source data, methodologies, and personnel have evolved since the first campaign, successfully maintaining alignment with the original design objectives established 35 years ago.

As shown in **Table 2.7**, when comparing CLC workflow–based production to the integrated CLMS data approach, the former is better positioned to support sustained ecosystem products, as it relies on a more robust framework for managing factors that may change over time.

Table 2.7. Continuity comparison

Legend of colours	
(green)	Meet the requirements
(yellow)	Partially meets
(pink)	Does not meet

	Current CLC Workflow with Incremental Improvements	Integrated CLMS Data Workflow
Continuity	The CLC workflow has demonstrated its long-term reliability despite numerous changes since the initial campaign, particularly in source data, methodologies, and involved experts, while maintaining alignment with its original design objectives.	A multiple-input approach may be subject to potential evolutions that cause the process to diverge from the original goal. This approach requires well-defined production criteria to ensure that the process remains aligned with the original goal.

2.2.7 Comparative criteria: synthesis of parameter evaluation

This section aims to acknowledge the strengths and limitations of both approaches and highlight the opportunity to develop a well-balanced solution. By carefully considering the unique strengths of each approach, it is possible to create a successful workflow in which methodologies complement and enhance one another. **Table 2.8** summarises how each approach addresses the requirements. The effectiveness of their implementation largely depends on available resources, IT capabilities, and the level of inter-institutional collaboration within each country.

Table 2.8. Summary of parameters evaluation

Legend of colours	
(green)	Meet the requirements
(yellow)	Partially meets
(pink)	Does not meet

Requirement parameter	Current CLC Workflow with Incremental Improvements	Integrated CLMS Data Workflow
Thematic content	Minor gaps in l2 for terrestrial and l1&l2 gaps for marine	Some gaps in L1 and gaps in L2
Temporal coverage	It could align with the requirements	It could align with the requirements

Temporal reference	It could align with the requirements	It could align with the requirements
Spatial resolution (MMU)	Non-compliant	Non-compliant, even if many inputs meet the requirements
Temporal frequency	Non-compliant	Non-compliant, even if many inputs meet with the requirements
Timeliness	Compliant. The production timeline could align with the required 18 months	Partially compliant
Completeness	MECE criteria are fully ensured within the CLC production and nomenclature	Achieving the MECE criteria requires effort due to the multiple diverse inputs
Continuity	The long time series of the CLC proved to have in place a good mechanism to ensure continuity	High dependency on different inputs of the production chain
Accuracy	Depends on the accuracy of the data capturing and processing at the local level	Depends on the accuracy of the individual CLMS products and on how they are integrated into the processing flow.
Consistency	The CLC proposes an approach to support consistency over time	Besides the CLC input, other CLMS data offer limited capacity to detect real changes over time
Transparency	A clear methodology is available	A clear methodology description is available for each product

The CLC-based workflow (option 1) continues to offer strong thematic coverage and enduring coherence, underpinned by its three-level hierarchical nomenclature and robust application of the MECE (Mutually Exclusive and Collectively Exhaustive) principle. This structure supports both completeness and high compatibility across reporting periods, as the process benefits from long-term consistency and harmonised methodologies. A major advantage of this approach lies in the systematic integration of local expertise and national knowledge, facilitating nuanced classification and ensuring the mapping process properly reflects on-the-ground realities across diverse European landscapes. This participatory model not only addresses regional differences in ecosystem characterisation but also enhances the credibility and acceptance of the resulting data.

However, operational challenges remain—notably the need to accelerate production cycles and improve spatial detail, while ensuring temporal consistency under changing retrieval methodologies. Existing limitations in timeliness and minimum mapping unit (MMU) must be addressed to meet evolving policy and user demands. Key priorities include reducing the update frequency from six to three years and refining the spatial resolution toward a 1–10 ha MMU standard.

Conversely, the integration-based approach (option 2), which leverages diverse ready-to-use data from the CLMS portfolio, demonstrates strengths in spatial resolution and update frequency. This approach can deliver more detailed and timely products in specific domains. Yet, at the current stage, it falls short in thematic completeness and nomenclature alignment, partly due to inconsistent coverage and varying update cycles among data sources. The primary hurdles relate to achieving comparability, interoperability, and continuity—all essential for supporting a harmonised, pan-European mapping solution. Realising these goals would require the creation of a centralised governance structure to coordinate disparate workflows, timelines, and quality standards.

Importantly, the two approaches exhibit complementary strengths and weaknesses. Option 1 ensures reliability, comparability over time, and a tested operational foundation, while option 2 can drive advances in spatial detail and frequency. **Table** visually illustrates these contrasts, highlighting, for example, the superior consistency of the CLC-based approach compared to the challenges faced by the integration-based workflow.

Yet, it is clear that neither approach alone fully meets the stringent requirements for an EU ecosystem extent product. Even if their best elements are combined, certain information and operational gaps would persist. Addressing these residual challenges will demand further innovation and dedicated resourcing, particularly in data harmonisation, integration methods, and update processes.

Overall, the CLC-based workflow remains the most viable and strategic base for future development: it is well-established, scalable, and designed for incremental enhancement. Its centralised setup reduces implementation complexity and initial investment, while retaining the flexibility to systematically incorporate new technologies and products from the integration approach over time. This foundation allows for continuous improvement without sacrificing methodological integrity or user trust.

Key questions moving forward include:

- Are current advances in AI, image analysis, and process automation mature enough to bridge the operational gaps—enabling faster, more detailed, and automated CLC production cycles?
- Can further integration of CLMS datasets effectively supplement the CLC workflow, particularly for ecosystem classes underrepresented or missing in the current nomenclature?
- What governance mechanisms are necessary to ensure ongoing alignment, comparability, and country-level endorsement as the mapping framework evolves?

A strategic combination of centralised CLC-based production—progressively adapted with input from artificial intelligence, expanded in-situ validation, and selective integration of high-value ready-to-use CLMS products—offers the most coherent path toward a robust, fit-for-purpose, and future-proof solution for EU ecosystem extent mapping. The rationale, methodology, and implementation steps for this tailored proposal are fully developed and detailed in Chapter 4 of this report.

3 Best practices and lessons from Pan-European and national mapping projects

This chapter reviews best practices and lessons learned from key European and international initiatives in ecosystem mapping. The insights gathered here directly inform and underpin the development of the tailored proposal presented in Chapter 4. By critically analysing operational approaches, technical innovations, and implementation challenges encountered in related projects, this section identifies practical solutions and methodological enhancements. These findings serve as an essential foundation for shaping an effective, coherent, and future-ready workflow—a workflow that will be systematically integrated, adapted, and refined within the proposal detailed in the next chapter.

3.1 Country experience from the current CLC workflow

3.1.1 Use of human abstraction capacity, local knowledge, and ancillary data

While centrally provided satellite imagery from the EEA is the primary source of information for updating (revising) CLC status maps and mapping CLC changes, there are other critical information sources that support implementation. Multi-temporal satellite imagery generally provides sufficient data on the biophysical characteristics of the Earth's surface—such as colour, spatial features, and temporal patterns—to identify land cover. However, land cover is just one type of information represented in the CLC nomenclature. It also includes data on land use (e.g., the level-1 class 'Artificial surfaces'), crop types (level-2 classes 21 and 22), management practices (classes 212 and 231 vs. 321), habitat types (e.g., 322, 323, 411, 412, 421), status indicators (classes 133, 324, 334), geographical information (classes 521, 522), and spatial patterns (level-2 classes 24, and class 313)—all of which require additional information for proper identification.

The human capacity for abstraction and synthesis is essential for interpreting these additional data: recognising patterns, heterogeneity, spatial and geographical context, and drawing on economic, social, and historical background knowledge—often referred to as local knowledge—is frequently necessary to distinguish between these classes. The importance of visual recognition becomes even more apparent when identifying and understanding changes. For example, while tree cover loss is easily detected through EO methods, it may result from various underlying processes—each reflecting different natural or economic trends and represented by distinct CLC change code pairs. Tree cover loss can result from forest management or reforestation (311-324), damage from wildfires (311-334), forest removal for construction (311-133) or mineral extraction (311-131), clearing for agriculture (311-221), or even the removal of a tree group within a golf course (which results in no change in CLC). Differentiating these processes is often only possible through photointerpretation based on the visual comparison of satellite images from two reference dates.

For these reasons, visual evaluation remains a critical tool for accurately mapping real changes in CLC and is necessary to complement semi-automated methods used to update CLC status layers.

Additionally, countries support the visual photointerpretation process with available national in-situ data, commonly referred to as ancillary data in CLC terminology. Frequently used ancillary data include digital orthophotos, topographic maps, LPIS data, building registers, cadastral data, National Forest Inventories, habitat maps, and sectoral thematic datasets (such as mines, irrigation, and wildfire records). These data are not directly integrated into the CLC, but instead are utilised through the photointerpretation process, enabling the 'absorption' of valuable, reliable, and often otherwise inaccessible (or prohibitively expensive) national information into the freely available European CLC product.

Many of the gaps identified in the previous chapter can be addressed by complementing automated methods with visual/manual photointerpretation or with visual review and approval of automatically detected features. Involving national experts in the workflow also facilitates the extraction of information that may only be available in national datasets, which are often inaccessible to third parties. However, accelerating administrative and governance processes remains necessary to ensure timely completion.

3.1.2 Semi-automated methodologies

While computer-aided photointerpretation (CAPI) has remained the predominant approach for creating new CLC inventories in most countries, several participating teams have transitioned to, or integrated, semi-automated methodologies within their production workflows to accelerate mapping and reduce manpower requirements. Most of these methodological innovations facilitate the production and updating of CLC status maps, although change mapping still largely relies on visual or manual methods.

The following chapter describes existing semi-automated approaches applied to the CLC update process as adopted by various countries. Identifying and applying these semi-automated methods can help make the monitoring process more agile for specific ecosystem types, reducing the need for extensive human intervention. These examples may benefit implementation in other countries and contribute to improved data frequency.

Since the CLC2006 inventory, some countries have begun applying semi-automated solutions—either fully or partially—to create CLC status layers and map CLC changes. Several of these solutions have proven successful, enabling operational and routine application. The primary objectives of these approaches are to reduce labour-intensive photointerpretation, increase accuracy and reproducibility, and improve compatibility with national databases. These solutions typically combine existing national GIS thematic datasets, satellite image processing technologies, on-screen digitisation (visual photointerpretation), and GIS-based generalisation. Nevertheless, some level of visual input or control is always incorporated, as human abstraction and interpretation cannot be entirely replaced by automated rules—particularly in the case of change mapping.

The implementation of these semi-automated, bottom-up approaches is heavily dependent on the types of LU/LC data available in each country. The availability of land use data is especially critical, as the functional use of land cannot be automatically derived from satellite data. Brief overviews of the methodologies used to date, based on the CORINE Land Cover Product User Manual (Version 1.0, p. 45), are presented below by country.

Three countries generate CLC data from a single high-resolution national LC/LU dataset, using advanced generalisation methods:

- **Germany:** The Digital Land Cover Model (DLM-DE) with a 1 ha MMU integrates national topographic reference data (ATKIS-DLM) and thematic remote sensing data, addressing national requirements. Since CLC2012, DLM-DE has supported European land monitoring by deriving high-resolution CLC data (1 ha MMU) and generalising it to the standard CLC MMU (25 ha). CLC-Change is generated by comparing DLM-DEnew and DLM-DEold and generalising the results to 5 ha. Both the status and change layers include manual editing and visual quality control (Federal Environment Agency Germany, 2012).
- **Portugal:** Beginning with CLC2018, Portugal has based its approach on higher-resolution national LC/LU data. An extensive thematic correction of CLC2012 was first conducted using the National Land Use and Land Cover Map (COS2010). COS, updated every five years, is fully compatible with CLC at level-3 and contains more thematic detail at levels 4 and 5, with a total of 225 classes and a higher geometric resolution (1 ha/20 m). After COS2010 was completed, a cross-comparison with CLC2012 was carried out, followed by manual correction of non-compliant objects larger than 25 ha. Corrections affected 10% of CLC2012 and 3% of COS2010. The corrected CLC2012 served as the basis for traditional photointerpretation-based change mapping in CLC2018. Additionally, >5 ha changes identified from the COS 2010–2015 change dataset was provided to photo-interpreters as ancillary information. COS change data are also produced directly, applying the established 'change mapping first' approach from CLC. However, interpreters had to account for the two-year date difference between CLC and COS surveys when creating CLC-Change2012-2018 (ETC-ULS, 2018).
- **Spain:** Since the CLC2012 inventory, Spain has employed a bottom-up approach using the National Information System on Land Cover & Use (SIOSE), a GIS at a 1:25,000 cartographic scale comprising polygons with minimum areas from 2 ha to 0.5 ha, depending on the LC class. SIOSE polygons individually describe land parcels by allocating sets of real landscape objects with associated percentages. The SIOSE data model includes 45 classes of landscape objects (e.g., buildings, bridges), 40 LC classes, and over 20 attributes to further characterise land cover. Its object-

oriented design enables seamless semantic generalisation for producing both the CLC status and change layers at the European scale (Hazeu et al., 2016).

Other countries create CLC by combining multiple existing national datasets and employing diverse EO data classification solutions:

- **Finland:** Applies advanced pre-processing of satellite images (using physical models such as atmospheric correction), classification of stratified imagery, regression estimation for forestry parameters, visual interpretation, integration of national land use data, and generalisation of high-resolution national data to European CLC standards. For national use, Finland employs level-4 CLC nomenclature at a pixel resolution of 25 m, providing greater thematic detail for forests, semi-natural areas, and wetlands (Finnish Environment Institute, 2005).
- **Iceland:** Thematic data are supplied by relevant national authorities and institutions, with several classes derived from remotely sensed data. The process involves GIS harmonisation and generalisation (National Land Survey of Iceland, 2009).
- **Ireland:** As part of CLC2012, Ireland conducted an extensive revision of CLC2006. A combination of ancillary vector data and satellite imagery interpretation was used to revise the 2006 dataset and identify changes between 2006 and 2012. LPIS was used for agricultural land use, vector data from the Forest Service assisted in mapping forests, and national water body data improved the delineation of water features and coastline. Ancillary data were generalised into a 5-ha intermediate dataset, coded to CLC classes, and intersected with CLC2006 to sort disagreements into change or revision categories using spectral interpretation in an object-oriented remote sensing platform. Areas not covered by ancillary data were assessed spectrally for misclassification. Manual photointerpretation was employed to map CLC changes not captured by automated methods and to verify the outputs (Lydon and Smith, 2014).
- **Norway:** Applies generalisation and merges data from land resource maps and national registers (Aune-Lundberg and Strand, 2010).
- **Sweden:** Uses theme-specific classification of satellite imagery (through interactive thresholding or automatic classification), visual interpretation, forest classification calibrated by National Forest Inventory data, and generalisation of high-resolution national data to European CLC standards. Sweden has produced 58 level-6 classes for national purposes with pixel resolution (25 m) and MMU values of 1, 2, 5, or 25 ha (Engberg, 2005).

3.1.3 Critical land use challenges

CLC national teams can offer valuable insights from their experience, particularly regarding the land use classes that are most challenging to map and may require specific ancillary data and expert knowledge. The latter involves interpreting satellite imagery within the local context to extract the necessary information. These examples can help guide efforts towards developing improved methods or resources to address critical aspects of land use mapping more efficiently.

3.2 Advancements in EO-based mapping

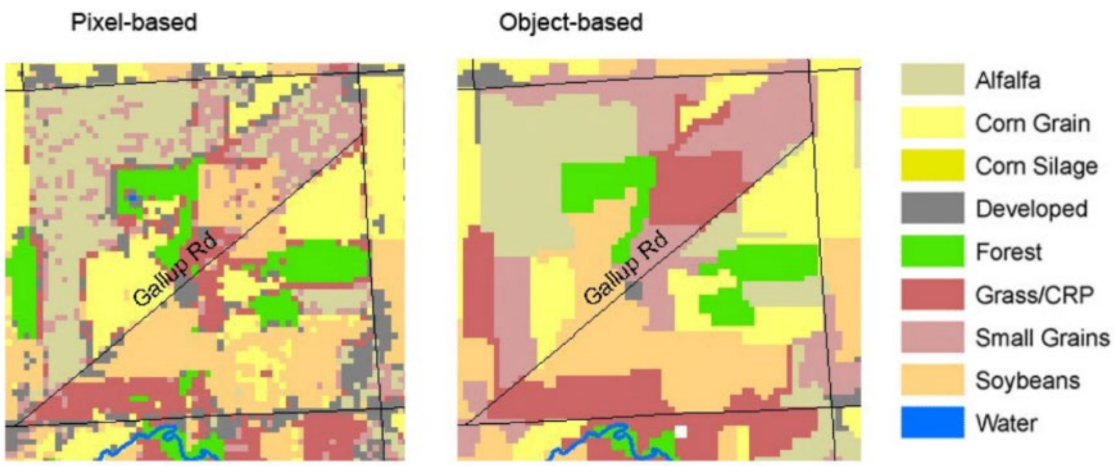
This section examines how recent technological advancements—including cloud platforms, improved infrastructure, and established methods and tools—can enhance the effectiveness and efficiency of the two proposed approaches. The focus is on both improving processing capabilities and facilitating the identification of specific land use gaps. Particular attention is given to ecosystem types and critical land use classes identified in the previous section.

3.2.1 Object-oriented approach

The object-oriented approach supports the detection of linear features and facilitates assessment of their connections and associations with neighbouring elements, within the context of an ecosystem's functional unit.

In land cover mapping, the object-oriented approach leverages advanced segmentation techniques to classify imagery. Unlike traditional pixel-based methods, which classify each pixel independently, Object-Based Image Analysis (OBIA) groups pixels into meaningful image objects using properties such as spectral, textural, geometrical, and contextual information (**Figure 3.1**). Similarly, Convolutional Neural Networks (CNNs) incorporate contextual information by analysing not just individual pixels but also their surroundings. This representation more closely aligns with real-world geographic entities than single-pixel approaches (Duro et al., 2012). OBIA effectively reduces within-class spectral variability and incorporates spatial relationships, resulting in more accurate and ecologically meaningful classifications. Consequently, this approach enhances classification accuracy, particularly in complex and heterogeneous landscapes (Kucharczyk et al., 2020).

Figure 3.1. Simple comparison of OBIA with traditional pixel-based approaches showing benefits of object-based mapping in a heterogeneous landscape (Powell & Colin, 2008)



Object-Based Image Analysis (OBIA) offers distinct advantages, particularly for the detection and delineation of linear features. This makes OBIA highly relevant in areas with complex spatial patterns, such as urban settlements, transportation infrastructure, and water bodies. By leveraging the shape and contextual properties of segmented objects, OBIA effectively differentiates linear and infrastructural features that may share similar spectral characteristics but have distinct geometries or spatial contexts.

OBIA has been thoroughly evaluated and validated in numerous scientific studies, frequently in direct comparison to traditional pixel-based classification methods. For example, Ye et al. (2018) systematically reviewed 209 journal articles published between 2003 and 2017 that assessed OBIA accuracy. They identified three critical stages in the process: sampling design, response design, and accuracy analysis. Their findings indicated an average accuracy of 85% across the reviewed studies, highlighting OBIA's robust performance across diverse applications. The study also noted the rising use of polygon-based accuracy assessments and emphasised important methodological considerations such as effective sampling of polygons, management of mixed polygons, and assessment of segmentation quality. Addressing these methodological challenges is essential to fully realise the benefits of OBIA.

Recent comparative studies further underscore the advantages of OBIA over pixel-based approaches. For example, Lei et al. (2024) reported significantly higher accuracy using multi-scale OBIA (91%) than single-scale OBIA (84%) or pixel-based methods (73%) for residential dwelling identification. Similarly, Dobrinić et al. (2024), using multitemporal PlanetScope imagery, found OBIA to deliver greater overall accuracy (85.43%) compared to pixel-based approaches (81.90%), while also reducing the “salt-and-pepper” effect commonly observed in pixel-based classifications. The enhanced performance of OBIA was especially apparent in classifying built-up areas and delineating agricultural fields. These studies also point out the importance of careful segmentation parameter selection to optimise outcomes.

In ecological applications, OBIA improves thematic accuracy by supporting the identification and classification of mixed or transitional ecosystems—such as ecotones or semi-natural habitats—through grouping pixels into ecologically meaningful units. While OBIA primarily relies on visual characteristics directly extracted from satellite or aerial images, it also allows for the integration of ancillary data (e.g., digital elevation models, climate data) to further refine classification results. This ensures higher thematic purity and minimises fragmentation in land cover maps.

For ecosystem extent data products—such as those required for ecosystem accounting frameworks like SEEA-EA—OBIA provides crucial benefits. Its capacity to generate spatial units closely aligned with ecological boundaries satisfies requirements for mutually exclusive and collectively exhaustive mapping (the MECE principle). OBIA also enables aggregating smaller segments into larger units while preserving typological relationships across segmentation levels. Its robustness in integrating multiple datasets further enhances the precision and thematic accuracy necessary for ecosystem monitoring and accounting, thus supporting accurate and consistent ecosystem extent analyses over time.
However, maximising the benefits of OBIA requires meticulous selection and optimisation of segmentation parameters. It is vital to balance over-segmentation, which can lead to unnecessary fragmentation, against under-segmentation, which may obscure critical ecological details. Other key challenges include sensitivity to input parameters, increased computational demand, and the complexity of developing suitable rulesets.

3.2.2 CLCplus, and other platforms

The CLCplus concept complements and extends the CLMS portfolio to better address growing user demands for pan-European land cover and land use monitoring, as well as related policy requirements.

The CLCplus framework[6], built around the CLCplus Core, serves as a centralised hub for harmonising diverse input datasets. However, it relies heavily on expert knowledge to ingest and harmonise data through a process of 'thematic translation.' As a result, outcomes may vary depending on the user's perspective or experience. To ensure consistency, the CLCplus Core tool would still require a centrally defined, harmonised decision-making framework agreed upon by various stakeholders.

Additionally, the framework has not yet been exploited to its full potential as originally envisioned. In particular, national data have not been made available beyond limited initial testing, leaving its added value in this context still uncertain.

One component of CLCplus that has been regularly implemented as part of CLMS production is the CLCplus Backbone: a high-resolution geospatial baseline dataset that provides harmonised land cover information across the entire European territory. However, the CLCplus Backbone product exclusively describes the characteristics (biotic or abiotic) of the material substrate at the Earth's surface, using a two-dimensional grid with a cell size of 10 meters. By applying thresholds based on the majority rule (assigning each grid cell to its most prevalent biotic or abiotic property), the product labels each pixel according to its dominant "land cover class" based on a generic nomenclature. Nevertheless, this product does not constitute a

[6] CLCplus: a framework for harmonized and tailored land cover information of Europe — Copernicus Land Monitoring Service

thematic dataset with spatial units that represent true physical features and thus is not suitable as a basis for defining ecosystem assets.

Despite its complexity, the CLC+ platform can function as a valuable integration tool, but it cannot fully overcome intrinsic limitations present in the input data itself.

Within the context of the proposed approaches, CLCplus could facilitate the integration of multiple input sources, including those from CLMS or other providers. Nonetheless, it is unlikely to fully resolve key challenges such as consistency and comparability, as discussed in Section 3.

3.3 Key projects and initiatives

This section explores how existing initiatives with similar objectives—specifically those mapping EU ecosystems for monitoring purposes—can inform improvements to the two proposed approaches. Insights from these initiatives can help make the approaches more effective and efficient, both in terms of processing and in supporting the identification of specific land use gaps.

3.3.1 EU Grassland Watch

The objective of the EU Grassland Watch (EUGW) project is to provide harmonised, long-term information about changes in grasslands in protected areas (Natura 2000). Grasslands are a key part of many landscapes, whether as hay meadows, grazing marshes or woodland clearings. They maintain biodiversity and food production but also influence ecological processes on a wide range of scales. These include critical functions such as pollination, water supply, carbon sequestration, and climate regulation.

Earth Observation is an obvious way to monitor grasslands, but their dynamic nature, subtle variations and unique management practices mean that conventional land cover mapping every few years rarely provides enough information.

Addressing these challenges, EUGW has successfully developed an operational approach that characterises grasslands not only according to their EUNIS Level 2 type, thus fulfilling classification needs, but also extends this by describing their management and productivity using modular components. This approach ensures compatibility with pre-defined nomenclature requirements while allowing a richer ecological description. Moreover, this characterisation is done on an annual basis, ranging from 1994 (the time of the designation of the first Natura 2000 sites) until "today".

The main challenges associated with the operational approach encompass:

- The wide range of grasslands, from urban to agriculturally used to natural types, and the absence of a universally agreed EO-based reference dataset for identifying them, poses a challenge. While no single authoritative mask exists, EUGW incorporates several grassland masks derived from CLMS products, alongside its own grassland mask. Each of the different CLMS products offers its own solution. EUGW therefore incorporates different grassland masks based on existing CLMS products as well as based on its own grassland mask. To explain further, the used grassland mask serves as a spatial filter that defines the area (delineation) where all subsequent characterisations, including type, management, and productivity, are performed. However, the important factor is supporting multiple mask options (based on HRL GRA, Land Cover, or internal derivations). Thanks to this, EUGW provides analytical flexibility to accommodate varying user needs and ecological interpretations. Thus, the user can then choose which mask to apply as the basis for their analyses.
- The most challenging and most important differentiation pertains not to mapping grassland extent, but to characterising their condition, particularly the intensity of their use. The problem can also

be reversed by considering the grassland condition as a discriminatory factor. As no single EO-based product can directly infer grassland management intensity, this aspect is inferred from proxies such as mowing or ploughing indicators.
- The final challenge is the temporal harmonisation of the time series, ensuring that only real changes are recorded in the data and transferred to the indicators shared with the end user.

Furthermore, the technical experiences and lessons learned within the COP4N2K project (predecessor of EUGW) offer some valuable insights. These experiences are mainly related to applying supervised machine learning algorithms, particularly Support Vector Machines (SVM), while the EUGW project uses random forest classification directly implemented on the OpenEO platform.

The methodological framework of COP4N2K applied SVM classification as a core component in both land use/land cover (LULC) mapping and in assessing grassland management intensity. SVM was used to classify multitemporal composites of Sentinel-2, Landsat, and Sentinel-1 (SAR) data, distinguishing key land cover classes and producing annual LULC maps between 1994 and 2018. Additionally, SVM was applied in vector-based classification to differentiate intensively and extensively managed grasslands. The performance and accuracy of the classifier were highly dependent on the availability, thematic reliability, and spatial representativeness of the reference datasets derived from CLMS N2000 mapping (2006/2012/2018). The project identified several limitations related to the quality of these training datasets (particularly inconsistencies and thematic inaccuracies in the reference labels), which led to misclassification risks and reduced model robustness in both LULC and management assessments.

One of the key insights from the COP4N2K project concerns the use of the polygons of CLMS Priority Area Monitoring product on protected areas in Natura 2000 (N2K), as in-situ reference data for training and validating the classification model. While the N2K 2006/2012/2018 datasets provided a valuable starting point, their thematic and spatial suitability proved problematic in many cases. Visual checks against very high-resolution (VHR) imagery revealed that numerous N2K polygons contained incorrect or outdated land cover labels. Furthermore, many polygons were too large and spatially heterogeneous, encompassing multiple land cover types. This led to inconsistencies during the training of the SVM classifier, as pixels within a single polygon could correspond to different real-world surfaces, violating the homogeneity assumption typically required for supervised classification. These issues were especially critical in the grassland management intensity assessment, where the thematic accuracy of reference labels was found to be inconsistent or, in some cases, almost random. The project highlighted the need for more refined selection criteria when using existing datasets as reference sources, ideally supported by independent visual interpretation or field validation.

In addition to the challenges with N2K datasets, the COP4N2K methodology recognised the broader conceptual distinction between classification and characterisation. Rather than relying solely on rigid categorical labels, the project introduced value-added products such as grassland phenology, mowing frequency, and biophysical traits to enrich the ecological characterisation of grasslands. These added-value products specifically include:

- **Grassland biophysics (BIO**): provides quantitative estimation of three basic biophysical variables of vegetation: leaf area index (LAI), leaf chlorophyll content (LCC) and leaf water content (LWC) retrieved from a time series of Sentinel-2 data. The output values are aggregated to the level of grassland polygons (in spatial domain) and are available at a monthly frequency (temporal domain).
- **Grassland phenology & productivity (PHENO):** provides estimation of 14 phenological variables, including, for example, start of season (SOS), end of season (EOS), length of season (LOS) and others, together with 6 variables related to grassland productivity. All of the considered variables are extracted from temporal profiles of Normalised Difference Vegetation Index (NDVI) derived from a series of Sentinel-2 data.

- **Grassland mowing (MOW):** provides information on grassland mowing intensity based on analysing SAR backscatter and SAR coherence (derived from Sentinel-1 data) together with a time series of NDVI (extracted from Sentinel-2 and Landsat data). The product informs about the occurrence of possible mowing events in grassland polygons with a monthly frequency.
- **Grassland ploughing & degradation (PLOUGH):** provides an indication of possible ploughing events or significant grassland degradation by detecting local anomalies with the sudden occurrence of bare soil within grassland areas. Detection of base soil occurrence is based on NDVI thresholding combined with analysing the differences of NDVI values in the monitoring period against the NDVI maximum corresponding to the last vegetation climax.

This shift acknowledges that nuanced ecological conditions (e.g., management intensity) may be better understood as gradients or multi-dimensional profiles rather than fixed classes. To support this approach, COP4N2K utilised a typology of in-situ data sources, including VHR-based visual interpretation, internal photo checks, and a carefully selected subset of "pure" N2K polygons. These served not only for model training but also for targeted validation and plausibility checks. The project's experience underscores the value of integrating multiple layers of in-situ evidence, especially when classifying dynamic processes such as mowing or management intensity that are difficult to observe directly through remote sensing alone.

Regarding the used input satellite data, the COP4N2K project relied on multitemporal composites generated from Sentinel-2 optical imagery, Landsat data, and Sentinel-1 SAR backscatter. Therefore, both COP4N2K and EUGW used the same types of Earth Observation data sources. The critical difference lies in how these data were integrated into the classification process. In COP4N2K, a single core product was created first, followed by several additional products, which did not influence the classification outcome of the core product. In contrast, EUGW employs a more integrated approach where all derived products directly contribute to the final characterisation of grasslands, establishing a strong interdependence among all layers.

Another key distinction is the timing of data fusion. In COP4N2K, optical and radar datasets (i.e., the optical and SAR branches) were processed separately, and results were fused only at a later stage. In EUGW, however, both optical and radar data are combined from the outset of the classification workflow, enabling a more synergistic use of the input sources throughout the entire processing chain.

In COP4N2K, these composites provided dense temporal profiles capturing vegetation and structural changes across target grassland sites. However, optical imagery is inherently sensitive to cloud cover and other atmospheric conditions, which can cause gaps in the time series. To ensure data reliability, pixel-level quality masks based on Sentinel-2 Scene Classification Layers (SCL) and Landsat Quality Assurance (QA) layers were employed. These masks filtered out invalid pixels affected by clouds, shadows, snow, or sensor issues, thereby improving the consistency of the input data and the resulting classifications.

One of the most significant methodological challenges in COP4N2K was assessing grassland management intensity, particularly distinguishing between intensively and extensively managed grasslands. As highlighted in the project, such management intensity cannot be observed directly from satellite data. Instead, proxies such as NDVI time series, textural metrics (e.g., Angular Second Moment), and SAR backscatter statistics were used and aggregated at the polygon level. These variables then served as input for a separate SVM-based classification workflow applied to vector data. While this approach was promising, its accuracy was highly sensitive to the quality of the reference data. Inconsistencies in the management labels assigned within the N2K datasets often led to confusion during training and validation. These findings reinforce the need for robust field calibration and reliable in-situ data when attempting to classify subtle ecological gradients such as management intensity.

Although formal lessons learned from EU Grassland Watch (EUGW) are not yet available due to the ongoing status of the production workflow, several methodological improvements and conceptual shifts from its predecessor project COP4N2K can already be identified.

1) Firstly, a key structural change lies in how spatial units are treated. While COP4N2K aggregated input data early on to polygonal boundaries (e.g., CLMS Natura 2000), EUGW retains a raster-based processing workflow throughout the classification process, postponing any aggregation to the distribution stage. This change avoids early loss of within-polygon variability and allows end-users to flexibly derive statistics at different spatial scales or with custom geometries.
2) Secondly, the product design evolved. In COP4N2K, grassland classification was disconnected from auxiliary indicators, such as cover and "grassland type" which, were fixed in the main product, while other descriptors (e.g., phenology, biophysics, or ploughing) were treated as independent add-ons. In contrast, EUGW adopts a more integrated approach, where all derived products actively influence the final grassland categorisation. This reflects a more dynamic and ecologically informed framework.
3) Thirdly, the nomenclature of grassland categories was significantly revised. COP4N2K followed the MAES classification, which mixed land cover and management concepts in a way that introduced ambiguities and overlooked valid combinations of grassland types and management regimes. EUGW addresses this by separating the production into four independent classification components: (1) grassland extent (based on land cover), (2) grassland type (biological classification), (3) grassland management intensity, and (4) grassland productivity. These components can be freely combined, offering users a more comprehensive and modular description of grassland ecosystems from multiple perspectives.

Altogether, the mentioned advances in EUGW reflect a maturing approach to ecosystem monitoring, rooted mainly in the lessons of COP4N2K and aligned with the needs of operational ecosystem accounting.

3.3.2 European Wetland Map

The European Wetland Map [7] (EWM) is an initiative promoted by the Horizon Europe projects AlfaWetlands and WETHORIZONS. It was published as the most comprehensive and up-to-date wetland map for the European region, locating, assessing, and merging the latest geospatial data on the distribution and types of floodplains and wetlands – coastal, mineral and peatlands in a bottom-up approach.

The map covers EEA38+UK, and it is based on a compilation and harmonisation of several global, European, and national datasets at 1 arcsec resolution (approx. 30m), distributed both as polygon and raster layers by country. This exercise, however, comes at the cost of a significant generalisation of the nomenclature, reducing the number of wetland classes to general levels, with a total of 10 classes corresponding to **4 wetland groups**, which differ considerably in terms of the characteristics and reference dates of the input data, being the following:

- **Floodplains:** this group includes a flood risk class with an area that would be flooded during a 100-year flood, according to Dottori et al (2021)[8]the potential floodplain and the actual riparian zone, the latter two based on the delineation from the Copernicus Riparian Zones product.
- **Coastal wetlands:** it includes the classes tidal flat, estuaries, lagoon, salt marsh and salines according to the Copernicus Coastal Zones products, with additional data extracted from the Global Wetland Map[9] (only for the tidal flat, salines, and salt marsh classes)

[7] https://zenodo.org/records/14883520
[8] http://data.europa.eu/89h/1d128b6c-a4ee-4858-9e34-6210707f3c81
[9] https://essd.copernicus.org/articles/15/265/2023/

- **Peatlands:** the peatland class builds on the update of the European Peatland Map from Tanneberger et al. (2017)[10] using data from over 100 different national and regional datasets.
- **Other wetlands:** this is a general 'wetland' class, containing wetlands that could not be clearly categorised in the other groups, most likely indicating mineral wetlands. Main source is the EEA Extended Wetland Ecosystem (EWE) layer 2012, including the following classes: *8 Managed or grazed wet meadows or pastures*, *9 Natural seasonally or permanently wet grassland*, *10 Wet heaths*, *11 Riverine and fen scrub*, *13 Inland marshes* and *14 Open mires*. Further wetland data were extracted from OpenStreetMap, including reed beds, wet meadows, and inland marshes.

With a view to mapping and monitoring of European wetlands, the EWM is an interesting resource as ancillary data, both to support the production of new wetland maps or for validation purposes, especially in the case of peatlands, where a huge effort was made to verify data accuracy and to include new information produced by Alfa Wetlands. However, the scope of the product and the definition of wetlands used differ from previous work by the EEA and Copernicus, which may hinder its integration into new mapping activities (more details available in Annex III). In this regard, it is clear that there is a limitation of reliable, up-to-date data on wetland coverage. Remote sensing products mostly lack thematic detail and they usually present relevant errors or inconsistencies in the classification, making them insufficient to meet European requirements. Efforts to produce reliable wetlands distribution maps from EEA and CLMS data without relying on CLC information leave gaps because EEA38 coverage is incomplete and limited to the Copernicus PAM mapping area.

This challenge could be overcome by expanding PAM's area of interest to cover wetland ecosystems not currently considered. It is worth mentioning that coastal and river-related wetlands classes existing in the EWE layer are covered by between 70% and 90%, so inland wetlands are mainly omitted. Considering the currently available information on wetland coverage, it would not be difficult to identify the main wetland areas in Europe and define a specific mapping area for a potential new PAM product that ensures the provision of more detailed data for wetland ecosystems, both spatially and thematically.

The utilisation of the Coastal Zones and other Copernicus products as the Riparian Zones has demonstrated good reliability in detecting wetland for certain classes. This is confirmed by the successful integration of the 'Coastal wetland' and 'Other wetland' groups within the final EWE data layer. Based on this consideration, it can be concluded that, without the geographical constraints imposed by the specific focus of the Copernicus product (i.e. on coastal or riparian zones), the mapping of the remaining European territory could be realised by applying the same methodology accordingly. The Coastal Zones product builds on a mixed methodology of automatic classification routines and visual interpretation of VHR satellite data, implementing tailored Coastal Zone nomenclature. Among the input images, it relies on data from SPOT (5, 6, and 7), Pleiades and WordView missions which guarantee resolution between 1.5 and 4 m. Numerous additional reference and in-situ data sources serve as ancillary data including CLC 2012/2018, Urban Atlas 2012/2018, GIO HR Layers, DWH_MG2_CORE_01 Coverage 2 (RapidEye, 5m), national orthophoto WMS services, as well as satellite imagery from Google Earth and Bing Maps;

Alternatively, the existing delineation of wetlands could be used to train models for extending the mapping and detecting the remaining wetlands not covered by the Copernicus product.

CLC information, despite its lower temporal resolution, could ensure consistency in the classification in these areas, given that, in a spatial data context, wetlands are distinct areas characterised by specific biotic, ecological, and edaphic conditions, influenced by the management applied; all aspects that are efficiently captured by the CLC approach.

[10] https://www.mires-and-peat.net/api/v1/articles/128692-the-peatland-map-of-europe.pdf

Being highly dynamic systems, changes in surface water, vegetation, and bare soils do not usually mean changes in the LULC class but are rather the result of seasonal or temporal changes linked to water availability in a specific region, usually intra-annual changes. Therefore, tailor-made products that consider factors such as the hydroperiod, soil moisture or water salinity are essential for the accounting and mapping of these complex and rapidly changing ecosystems with the level of detail required.

3.3.3 SEPLA project

In 2021, the Joint Research Centre (JRC) launched the SEPLA (Satellite-based mapping and monitoring of European peatland and wetland for LULUCF and agriculture) project under the work programme with DG CLIMA and technical experts from the national administration and research bodies in 10 EU countries. The main objective of the project was to develop and agree on methods for comprehensive inventories of wetlands and peatlands that support the monitoring of their preservation and restoration. An emphasis was placed on the cross-cutting nature of the available spatial data of peatlands and on the diversity of application domains the data could originate from. The inventory of these areas would require a common perception and definition of what a wetland or peatland is, which has been proven to be challenging.

SEPLA elaborated a prototype and technical guidance for Earth observation-based monitoring of peatlands or wetlands. Methodologies for the semantic description of the land cover, land use and soil-related aspects were proposed as a precondition for this framework for peatland monitoring. Adapting feature-level Checks by Monitoring concepts to deal with ecosystems and with natural processes, the project elaborates examples of monitoring scenarios and tests several Sentinel-derived signals and markers. The peatland monitoring complexity is addressed by its decomposition into smaller and more manageable components. This involves a controlled breakdown of the requirements, feasibility and timing, the elimination of unnecessary processing and the prioritisation of queries towards specified expectations. The three key elements of the monitoring system are discussed: the observation method, the temporal phenomenon that reveals an attribute or characteristic and the spatial feature of interest that exhibits the observed phenomenon. Additional EO-based methods that rely on both passive and active satellite sensors are investigated. A proxy method based on multiannual farmers' declarations for handling non-monitorable land aspects is described.

The project also documented a summary of study cases that explore the possibilities of discrimination between the managed grassland on organic and on mineral soils, the in-depth characterisation of areas on organic soils and a machine learning based peatland mapping approach. All analyses were performed using data provided by the participating Member States: Bulgaria, Denmark, Ireland and Latvia combined with publicly accessible data and products provided by Copernicus Sentinel data access and distribution services and the Copernicus Land Monitoring Services. Despite their promising outcomes, these case studies revealed that ground data collection will be essential for learning, optimisation, and validation of the EO-based monitoring methods. Tailored regional-based approaches, relying on local contextual information and using EO-based solutions or Copernicus downstream services, can be expected to contribute to a workable solution.

The project concluded that the satellite-based monitoring of peatland and wetland is a very complex, yet feasible task, but it requires multidisciplinary expert knowledge and diverse types of data. It also requires the integration of satellite data with ground data, with hydrological models or with ecosystem models. The establishment of a community of practice will likely be essential for an effective uptake of novel technology and for the transfer of know-how. Ongoing projects that further explore these options are the GlobalWetlandWatch project (sponsored by Google, see https://www.globalwetlandwatch.org/) and the EO4WI (sponsored by ESA, see https://www.eo4wi.earth/).

3.3.4 PEOPLE-EA/ARIES FOR SEEA

Artificial Intelligence for Environment & Sustainability (ARIES), developed by researchers at the Basque Centre for Climate Change (BC3), is an open-source modelling platform, leveraging the power of artificial intelligence (AI) to advance research and innovation for environmental sustainability. Researchers worldwide can add their own data and models to web-based repositories.

Using ARIES technology, the ARIES for SEEA Explorer application allows users to compile ecosystem accounts for any user-specified area in the world and accounting period (from 1992 onwards), using accessible data layers and models, consistent with the SEEA Ecosystem Accounting. The ARIES for SEEA Explorer can be accessed in two different ways: using a web browser or using ARIES' k.LAB software.

The current ARIES for SEEA Explorer functionalities are restricted to assessing:

1. ecosystem extent (based on the IUCN Global Ecosystem Typology)
2. condition (for forest ecosystem types)
3. selected ecosystem services in physical and/or monetary units (crop provisioning, crop pollination, climate regulation, sediment regulation)

The ARIES for SEEA Explorer automates model selection based on a user's specific request. It chooses the "most appropriate" model for the location, spatiotemporal resolution and account specified and depending on the models and data sources accessible to the system. For example, when compiling extent accounts, the ARIES for SEEA Explorer uses CORINE land cover data for EU countries, but ESA-CCI (or GLC-FCS30D) for non-EU countries, unless there are national data made accessible by these countries. For this, it would be essential that all country data be validated by countries before being placed in the global database.

IUCN's Global Ecosystem Typology (GET) has been selected as the SEEA Ecosystem Type reference classification for the purposes of international reporting and comparison. The ARIES for SEEA Explorer currently enables the modelling of 29 "Level 2.5" ecosystem types that are a mix of Level 2 biomes and Level 3 EFGs of the IUCN's Global Ecosystem Typology. These include 21 terrestrial biomes and EFGs, seven wetland EFGs, plus the open water biome (**Table 3.1**).

The methods for mapping these "Level 2.5" ecosystem types are derived from temperature domains, landform and elevation data (according to the method from Sayre et al. 2020), with aridity domains from UNEP 1997, combined with land cover data in a lookup table (**Table 10**). This enables the mapping of ecosystem change over time using the best available data. The 10-degree threshold for warm-month temperature, proposed by Sayre et al. 2020, was relaxed to 14 degrees to enable mapping of alpine tundra in low-latitude regions.

Table 3.1. Mapping the 29 "Level 2.5" ecosystem types of the IUCN's Global Ecosystem Typology, derived from temperature domains, landform, elevation, aridity and land cover data

Landcover	Aridity	Mean temperature (C)	Landform	Elevation (m)	Ecosystem type
Forest	> 0.65	> 18	All but mountains	All	Tropical-subtropical lowland forest
Forest	0.05 - 0.65	> 18	All	All	Tropical-subtropical dry forest & scrub
Forest	> 0.65	> 18	Mountains	All	Tropical-subtropical montane rainforest
Forest	All	0 - 18	All	All	Temperate forest

Forest	All	< 0	All	All	Boreal & temperate montane forest & woodland
Shrubland	> 0.2	> 24	All	All	Seasonally dry tropical shrubland
Shrubland	> 0.2	10 - 24	All	All	Seasonally dry temperate heath & shrubland
Grassland/shrubland/savanna	> 0.2	< 10, and mean temperature of warmest month > 14	All	All	Cool temperate heathland
Sparse vegetation, bare area, lichen/moss	> 0.2	> 0	All	All	Rocky pavement, lava flow, & scree
Grassland	> 0.2	10 - 18	All	All	Temperate & subhumid grassland
Savanna	> 0.2	10 - 18	All	All	Temperate woodland
Grassland, savanna	> 0.2	> 18	All	All	Tropical & subtropical savanna
Grassland/shrubland/savanna, sparse vegetation, bare area, lichen/moss	0.03 - 0.2	0 - 10	All	All	Cool desert & semi-desert
Grassland/shrubland/savanna, sparse vegetation, bare area, lichen/moss	< 0.03	> 0	All	All	Hyperarid desert
Grassland/shrubland/savanna, sparse vegetation, bare area, lichen/moss	0.03 - 0.2	> 10	All	All	Other desert & semi-desert
Grassland/shrubland/savanna, sparse vegetation, lichen/moss	All	< 0, and mean temperature of warmest month < 14	All but mountains	All	Polar tundra & desert
Grassland/shrubland/savanna, sparse vegetation, lichen/moss	All	< 0, and mean temperature of warmest month < 14	Mountains	All	Alpine grassland & shrubland
Bare area	All	< 0, and mean temperature of warmest month < 14	All	All	Polar-alpine rocky outcrop
Glacier & perpetual snow	All	All	All	All	Ice sheet, glacier, & permanent snowfield
Agricultural land	All	All	All	All	Cropland
Artificial surfaces	All	All	All	All	Urban & industrial ecosystem
Inland swamp	All	> 24	All	All	Tropical flooded forest & peat forest
Inland swamp	All	10- 24	All	All	Subtropical & warm temperate forested wetland
Inland swamp	All	<10	All	All	Boreal & cool temperate palustrine wetland
Mangrove	All	All	All	All	Intertidal forest & shrubland
Wetland	> 0.65	> 10	All	> 5	Warm temperate & tropical marsh
Wetland	< 0.65	> 10	All	> 5	Episodic arid floodplain
Forest	< 0.05	> 18	All	All	Episodic arid floodplain

Inland swamp & Wetland	All	< 10	All	> 5	Boreal & cool temperate palustrine wetland
Wetland	All	All	All	< 5	Coastal salt marsh & reed bed
Water body	All	All	All	All	Aquatic

The Pioneering Earth Observation Applications for the Environment - Ecosystem Accounting (PEOPLE-EA) project, funded by the European Space Agency (ESA) and implemented between October 2022 and October 2024, studied the relevance of Earth observations for SEEA EA and explored its use for terrestrial and freshwater ecosystems through active involvement of five National Statistical Agencies in Europe, or their representatives, known as Early Adopters (Netherlands, Norway, Italy, Greece, Slovakia).

Based on the experience of the two technological project partners, VITO and the Basque Centre for Climate Change (BC3), the project developed a system-of-systems, the ARIES for People Explorer (ARIES for PEOPLE-EA). PEOPLE-EA integrates three state-of-the-art technologies:

1. ARIES, a software platform designed to integrate models via the use of well-defined scientific concepts and used as an Artificial Intelligence (AI) tool for rapid natural capital accounting
2. OpenEO, an open application programming interface (API) that connects to several clients (Python, R, JavaScript) to perform data operations on big Earth Observation data stored in the cloud
3. INCA project, which developed an ecosystem service models-based approach fully compliant with the SEEA EA to support European Member States in reporting ecosystem accounts

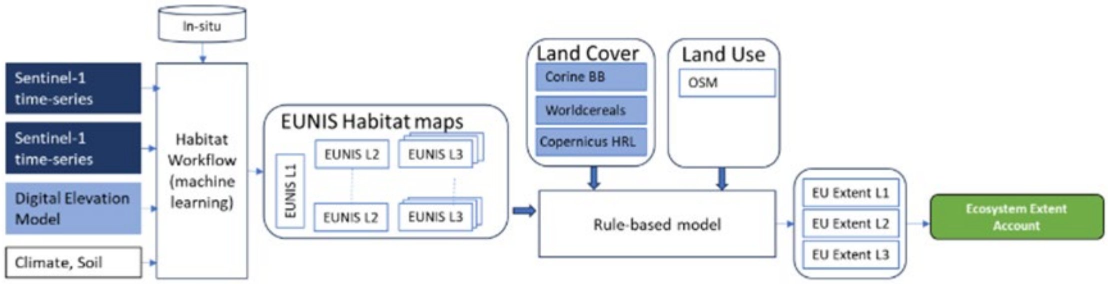
The use of Earth Observation for ecosystem accounts was demonstrated by the development of five workflows (EUNIS habitat mapping and ecosystem extent accounting, forest condition accounts, coastal condition, wood provision and nature-based tourism). The workflow for EUNIS habitat mapping and ecosystem extent accounting were explored in Slovakia and Greece and served as a basis for further development through the World Ecosystem Extent Dynamics (WEED) project.

For EUNIS habitat mapping and ecosystem extent accounting, the ecosystem types are classified according to the European ecosystem typology. Three different possible approaches were identified:

1. Derive from Corine Land Cover, complemented with additional Copernicus Land Monitoring (CLMS) datasets and optionally applying some national corrections. This approach is explored by several countries (e.g., Ireland) and by the European Environment Agency (EEA) and is defined as the **Top-Down** approach.
2. Derive from national dataset(s) and crosswalk(/model) to the EU typology, optionally gap-filled with public datasets (i.e., CLMS). This approach is pursued by several countries (e.g., the Netherlands, Hungary, and Germany) and is known as the **Bottom-up** approach.
3. Derive from vegetation maps and land cover characteristics, complemented with additional land use and other information, and is defined as the **Veg-map** approach.

For further exploration in the PEOPLE-EA project, the Veg-map approach was chosen. Its focus on natural vegetation classes based on EUNIS typology enables not only providing detailed (level-3) habitat maps but also its direct crosswalk to the EU extent typology at level-3. However, it first requires generating the EUNIS habitat maps, which were produced by combining geospatial datasets (Sentinel-1 and Sentinel-2 time series, Digital Elevation Model, climate, and soil) with ground-truth sampling points. The hierarchical EUNIS habitat maps were thereafter joined with Land Cover (Corine backbone, Copernicus High Resolution layers, WorldCereal) and Land Use (Open Street Map) through a rule-based model to generate Ecosystem extent maps at Level 1 up to Level 3 (**Figure 3.2**).

Figure 3.2. High-level workflow for ecosystem extent accounts



Ecosystem Extent accounts were generated for Slovakia (entire country) and Greece (Peloponnese) for the year 2020, with a focus on detailing forest & woodland and coastal ecosystem types up to Level 3. The extent account was validated by Slovakia for the Slovak Karst and Devon Lake areas by comparing it with ground-truth data assembled in 2024. At Level 1 and Level 2, an accuracy of above 90% was reached, while at Level 3, the accuracy dropped to slightly higher than 60% (Ecosystem Extent Account | PEOPLE-EA[11]). This drop was mostly attributed to the poor detection of Ravine and Carpinus and Quercus mesic deciduous forest types, as well as wetlands. For Greece, an evaluation was done based on expert judgement and visual interpretation, with a focus on well-known areas by the early adopters' experts.

The technical developments and results of PEOPLE-EA people led to the elaboration of the Evolution Roadmap document, which outlines key insights into leveraging Earth Observation (EO) data for mapping ecosystems across various typological levels. Critical lessons were learned regarding the effectiveness of integrating EO datasets with field reference data to produce scalable, wall-to-wall ecosystem datasets and the importance of user engagement and capacity building in the effective application of EO tools. The Evolution Roadmap document outlines **specific R&D topics** relevant to the different accounts, limited to terrestrial and freshwater ecosystems. **Table 3.2** below lists the key priorities, based on a ranking of priorities by stakeholders (incl. EEA) according to the pre-defined criteria.

Table 3.2. R&D priority topics for extent accounts

Extent account	Comments
Developing a protocol to map changes in ecosystem extent (assuring that only real changes are included in the accounts, not changes due to errors or biases in spatial data)	Changes in ecosystem extent are often more policy relevant than the areas themselves. At the same time, these trends may be unreal due to differences in datasets or methodologies, or due to uncertainties in the data. Such sources of change should be removed before statistically relevant trends are presented to the user.
Enhancing the use of EO data to monitor and analyse level 3 or other specific ecosystem types linked to Nature-Based Solutions. E.g., peatlands (area, location, drainage, subsidence)	Globally critically relevant, for ecosystem accounting as well as for LULUCF analysis and reporting to UNFCCC as well as for biodiversity strategy and climate change mitigation efforts. Incorporating high-resolution hyperspectral imaging and fully exploiting EO time-series can better differentiate detailed habitat types (e.g., forest types, clear cuts as grasslands, wetlands, etc.)
Mapping crop types, relevant for ecosystem services.	This important data layer is currently missing at the EU scale or limited to some crop types, even though technically, maps of crop type can be produced with EO data with high accuracy (>90% depending upon crop

[11] https://esa-people-ea.org/en/results/demonstrator-accounts/ecosystem-extent-account

	types, slopes, and field sizes). However, this is more challenging in the case of smallholder holdings and intercropping practices, and mapping crops for specific ecosystem services, such as pollination. The potential of the upcoming hyperspectral missions could be investigated to fill this gap.
Mapping linear landscape elements, and developing protocols on whether to include these (e.g., hedgerows) in extent and/or in condition accounts	Linear landscape elements are critical landscape elements for biodiversity and several ecosystem services such as air filtration, local climate regulation and pollination, and capture diversity within smallholder and agroforestry systems. Their presence or absence needs to be covered in ecosystem accounts, but challenges remain, given the diversity in width, length, and structure of these elements.
Dealing with the heterogeneity of urban and agricultural mixed landscapes.	Guidance and protocols need to be developed, based on best practices, on how to map fragmented landscapes, e.g., on the MMU to apply, and on how to aggregate classes. This is of relevance for spatially diverse urban and mixed agricultural landscapes. Where possible, localised high-resolution EO data should be used in support of wall-to-wall EO datasets.

3.3.5 BigEarthNet

BigEarthNet v2.0 (reBen) is a European land cover image archive built from Sentinel-1 and Sentinel-2 image patches. It is the world's first open library dedicated to Earth observation (EO) training data and is constructed to support deep learning (DL) studies for remote sensing image analysis.

BigEarthNet (https://bigearth.net/) consists of 549,488 pairs of Sentinel-1 and Sentinel-2 image patches with spectral bands at 10, 20, and 60-meter resolution. The satellite images were acquired in different seasons between June 2017 and May 2018, covering 10 countries (Austria, Belgium, Finland, Ireland, Kosovo, Lithuania, Luxembourg, Portugal, Serbia, and Switzerland).

Each image patch was associated with a pixel-level reference map and multiple land-cover class labels (i.e., multi-labels) that were derived from CORINE Land Cover 2018 (Sumbul et al. 2019). Most image patches contain at most five labels, based on the thematically most detailed Level 3 class nomenclature from CORINE Land Cover 2018. However, some CLC classes are challenging to accurately describe by only considering (single date) BigEarthNet images (Sumbul et al. 2021). An alternative nomenclature of just 19 classes was therefore introduced based on the properties of BigEarthNet-MM images (**Table 3.3**).

Table 3.3. Reference map class IDs associated to the CLC class nomenclature and 19 classes nomenclature

Class ID	CLC Level-3 Class Nomenclature	19 Classes Nomenclature
111	Continuous urban fabric	Urban fabric
112	Discontinuous urban fabric	Urban fabric
121	Industrial or commercial units	Industrial or commercial units
122	Road and rail networks and associated land	Unlabeled
123	Port areas	Unlabeled
124	Airports	Unlabeled
131	Mineral extraction sites	Unlabeled
132	Dump sites	Unlabeled
133	Construction sites	Unlabeled
141	Green urban areas	Unlabeled
142	Sport and leisure facilities	Unlabeled
211	Non-irrigated arable land	Arable land
212	Permanently irrigated land	Arable land
213	Rice fields	Arable land
221	Vineyards	Permanent crops
222	Fruit trees and berry plantations	Permanent crops
223	Olive groves	Permanent crops
231	Pastures	Pastures
241	Annual crops associated with permanent crops	Permanent crops
242	Complex cultivation patterns	Complex cultivation patterns
243	Land principally occupied by agriculture, with significant areas of natural vegetation	Land principally occupied by agriculture, with significant areas of natural vegetation
244	Agro-forestry areas	Agro-forestry areas
311	Broad-leaved forest	Broad-leaved forest
312	Coniferous forest	Coniferous forest
313	Mixed forest	Mixed forest
321	Natural grassland	Natural grassland and sparsely vegetated areas
322	Moors and heathland	Moors, heathland and sclerophyllous vegetation
323	Sclerophyllous vegetation	Moors, heathland and sclerophyllous vegetation
324	Transitional woodland/shrub	Transitional woodland, shrub
331	Beaches, dunes, sands	Beaches, dunes, sands
332	Bare rock	Unlabeled
333	Sparsely vegetated areas	Natural grassland and sparsely vegetated areas
334	Burnt areas	Unlabeled
335	Glaciers and perpetual snow	Unlabeled
411	Inland marshes	Inland wetlands
412	Peatbogs	Inland wetlands
421	Salt marshes	Coastal wetlands
422	Salines	Coastal wetlands
423	Intertidal flats	Unlabeled
511	Water courses	Inland waters
512	Water bodies	Inland waters
521	Coastal lagoons	Marine waters
522	Estuaries	Marine waters
523	Sea and ocean	Marine waters
999	Unlabeled	Unlabeled

Since its release, BigEarthNet has paved the way for the development of different benchmark remote sensing datasets. For instance, in the ben-ge dataset is constructed by adding geographical (e.g., climate-zone and topographic data) and environmental (e.g., weather and seasonal information) data to the BigEarthNet patches (Clasen et al. 2024). Another example is the RSVQAxBEN dataset, adding semantic content-related questions and answers formulated in natural language, which is presented in for visual question answering problems (Clasen et al. 2024).

3.3.6 Other relevant projects and initiatives

National Italian land cover and land use map

ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale) is Italy's national environmental agency, which recently has produced a highly detailed, nationwide land cover and land use (LCLU) map for 2023. The primary inputs for this dataset are freely available data products from the Copernicus Land Monitoring Service. Italian experts combine the most recent Corine Land Cover layer, the CLCplus Backbone raster layer, and all the Priority Area Monitoring products to produce a highly detailed, wall-to-wall dataset that covers the entire country of Italy. ISPRA experts consider this data set to be a valuable tool for monitoring compliance with the EU's new Nature Restoration Regulation, contributing to Italy's reporting obligations under the United Nations' Sustainable Development Goal, and playing a role in Italy's greenhouse gas inventory, specifically for LULUCF (Land Use, Land Use Change, and Forestry) emissions assessments.

Copernicus Global Land Cover and Tropical Forest Mapping and Monitoring Service

The new Copernicus Global Land Cover and Tropical Forest Mapping and Monitoring service (LCFM) will provide a dynamic global land cover service for the years 2020-2026. It will produce land surface categorisations and land cover features at regular intervals and compile them into global annual land cover maps and tropical forest monitoring products at 10m resolution. A consortium led by VITO with partners GAF, GeoVille, GFZ, IGN FI and IIASA and the support of two subcontractors, Sinergise and Telespazio Ibérica, has been awarded by the JRC to implement this new service.

The LCFM contract started in January 2024 with a ten-month ramp-up phase. During this phase, the innovative workflow will be tested, and the results will be refined using a subset covering around 10% of the total land area. Once the results of the ramp-up phase have been evaluated, full-scale production of the annual and sub-annual land cover and forest monitoring products for 2020 will begin. The first public results are expected to be available in early 2025. Subsequently, the service will expand its scope and produce similar products for the years 2021 to 2026.

CUP4SOIL project

CUP4SOIL (Copernicus user uptake for soil information products) aims at a downstream service to support national and European agencies for reporting on soil health and quality. This further underpins the pre-operational Soil Monitoring System currently being developed within the European Space Agency's (ESA) WorldSoils project with the potential to serve as a component of the Copernicus Land Monitoring Service. Using synergies between this action and the ESA WorldSoils project is streamlining the activities and boosting the user uptake.
CUP4SOIL generates European-wide soil information products based on Sentinel-1 and Sentinel-2 data. DLR is creating several soil-related input products, such as soil reflectance composites, information about the cover frequency of soils and the vegetation dynamics on a high spatial resolution (20 m). These data products flow into the high-performance computing environment of ISRIC that generates information about soil organic carbon content, texture pH values, etc. using digital soil mapping approaches. The data is being made publicly available.

Bare Surface Frequency product from the SoilSuite, has potential for detecting ecosystem types "Sparsely vegetated ecosystems" and "Coastal beaches, dunes and wetlands", as part of the EU-wide mapping of ecosystem extent.

Lifewatch
The European Infrastructure for Biodiversity and Ecosystem Research (Lifewatch) aims at providing data and analytical tools focused on pressing scientific issues related to biodiversity. A contribution to this

initiative is the description in spatial terms of the biotopes in Europe (EU-38 countries) on a yearly basis. A geographic database has been designed after identifying the specific needs of the biodiversity and macroecology communities. The methodology is organised around the concept of spatial regions, which are landscape units with similar properties. Spatial regions are delineated by the segmentation of filtered images and characterised using a set of remote sensing data and GIS-derived indices. Preliminary results with RapidEye images were promising, and new research is being planned with the Sentinel-2 mission.

Biodiversity Meets Data (BMD)
Biodiversity Meets Data (BMD) is an EU-funded initiative under Horizon Europe, which began in March 2025 and runs until February 2029. Its mission is to create a Single Access Point (SAP) for biodiversity data and monitoring tools across Europe — enabling natural-resource managers and policymakers to access high-throughput monitoring tools (including AI taxon identification), mobilise historical baseline data, explore spatial environmental datasets, and perform analyses in virtual research environments (VREs) covering terrestrial, freshwater and marine systems. By harmonising data and providing scalable analysis platforms, BMD aims to support evidence-based biodiversity conservation, improve reporting under EU nature directives, and help reverse biodiversity loss across Europe.

EAGLE
The EAGLE Group (Action Group on Land monitoring in Europe) is a collaborative network of land-monitoring experts that, since 2008, has developed a semantic and technical framework for harmonised land-cover/land-use information across Europe. Within CLMS, EAGLE is recognised as a core component supporting the shift from traditional “classification” of land cover to richer characterisation of the landscape — enabling integration of different datasets, sensors and methodologies under a consistent model. The EAGLE conceptual model underpins the next-generation product framework in CLMS (notably the CLCplus initiative) to ensure interoperability between national and European land-cover/land-use products through shared semantics, standardised data models, and integration workflows.

3.4 AI-based experiences

The integration of artificial intelligence into Earth Observation (EO) has transformed land cover, land use, and ecosystem mapping through two principal methodologies: machine learning and machine reasoning. Machine learning approaches identify patterns directly from satellite imagery and in-situ measurements, facilitating the classification of land cover types, detection of ecosystem changes, and quantification of biophysical parameters using supervised and unsupervised algorithms. Deep learning variants—particularly convolutional neural networks (CNNs)—have shown exceptional capability for extracting features from multi-spectral and multi-temporal satellite data without the need for explicit feature engineering.

Complementing this data-driven approach, machine reasoning leverages "digital twin" models that encode ecological knowledge and biophysical processes, supporting inference based on established theoretical frameworks rather than solely on statistical correlations. These reasoning systems can integrate expert knowledge on ecosystem dynamics, succession patterns, and biogeochemical cycles, enabling more robust interpretation of remote sensing data, particularly in data-sparse regions.

The synergistic application of both methodologies—combining the pattern recognition power of machine learning with the knowledge representation and inference capabilities of machine reasoning—provides a promising framework for comprehensive ecosystem characterisation. This integrated approach enhances our capacity to monitor biodiversity, quantify ecosystem services, and assess landscape-scale ecological processes across spatial and temporal scales.

3.4.1 Enhancing ecosystem mapping with AI technology for critical classes

For ET1 (cemeteries) and ET2 (temporal grasslands, mixed agricultural patterns), traditional pixel-based classification approaches often struggle due to heterogeneous spatial arrangements and pronounced seasonal variations. Machine learning methods that leverage temporal sequences of imagery—particularly those utilising multi-temporal image segmentation—could improve the detection of distinctive phenological patterns associated with these ecosystem types.

The classification challenges of ET4 (broadleaved evergreen forests, plantations) may be addressed using deep learning techniques that can recognise subtle textural and structural differences between natural and plantation forests. Incorporating LiDAR data to assess canopy structure characteristics could further enhance classification accuracy.

For ET6 and ET11 (differentiation between land/coastal ecosystems and artificial shorelines), the transitional nature of these boundaries poses significant challenges. Combining edge detection algorithms with contextual analysis has the potential to improve classification accuracy beyond the capabilities of traditional methods.

3.4.2 AI-enhanced land cover mapping projects

The following examples from European research initiatives were identified through web searches supported by LLM queries. They demonstrate that AI is already making significant contributions to ecosystem mapping across the continent. These applications show promise in addressing the limitations of traditional approaches like Corine Land Cover by improving spatial resolution, enabling more frequent updates, and enhancing the classification of complex or dynamic ecosystems.

Most projects leverage machine learning algorithms to process satellite imagery, with many incorporating additional data sources to improve accuracy. The multi-temporal capabilities of several initiatives directly address the need for monitoring seasonal and long-term changes in ecosystems. As these technologies mature and become more widely adopted, they have the potential to transform environmental monitoring, conservation planning, and land-use decision-making throughout Europe.

These examples confirm that while AI offers powerful new capabilities, its most effective application comes through complementing rather than replacing traditional ecological expertise. Apart from being very useful in depicting features and patterns from satellite data, AI models can effectively integrate other types of data (soil, landform, land use, and even socio-economic data) to act as covariates in the learning model applied.

The Ulmas-Walsh Map for Ireland.
One of the most notable applications of AI in European ecosystem mapping comes from Ireland, where researchers have developed the Ulmas-Walsh map. This innovative project employed Convolutional Neural Network (CNN) machine learning algorithms to segment Sentinel-2 satellite imagery into various land-cover classes (up to circa CLC level 2 classes). The resulting 10-meter resolution map significantly outperformed traditional methods in terms of accuracy, demonstrating AI's capacity to identify features not correctly labelled in the CORINE database.

The Ulmas-Walsh map addresses several limitations mentioned in the query, particularly regarding the detection of seasonal variations in land cover. The system can identify features that change throughout the year, such as Turloughs (seasonal lakes), flood plains, and rotational crops—precisely the kind of dynamic ecosystems where traditional CLC mapping struggles. This capability for on-demand updates allows for near real-time monitoring of land cover changes, limited only by cloud cover conditions.
Source: https://asr.copernicus.org/articles/18/65/2021/

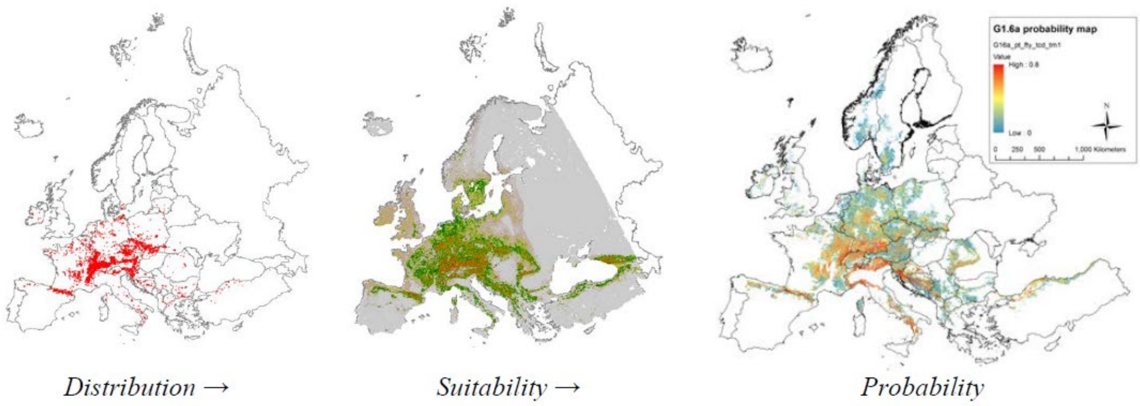
European Terrestrial Habitat Mapping Initiative

Through a series of contracts issued by the European Environment Agency (EEA) in the period 2012-2022, the terrestrial habitats of the EUNIS habitat classification were revised and linked to vegetation plot data (relevés) coming from the EVA database. The spatially explicit vegetation plots enabled EEA's European Topic Centre on Biological Diversity (ETC/BD) to prepare EUNIS habitat distribution maps and subsequently model high-resolution European suitability and probability maps employing AI/ML techniques. Neural Networks and Deep Learning methods were applied to environmental data and satellite-based datasets, including the Copernicus High Resolution Vegetation Phenology Product (HR-VPP).

Source: ETC/BD Technical paper 2/2019: Processing European habitats probability maps for EUNIS Forest (T), Heathland, scrub and tundra (S) and Grassland (R) habitat types based on vegetation relevés, environmental data and Copernicus land cover.

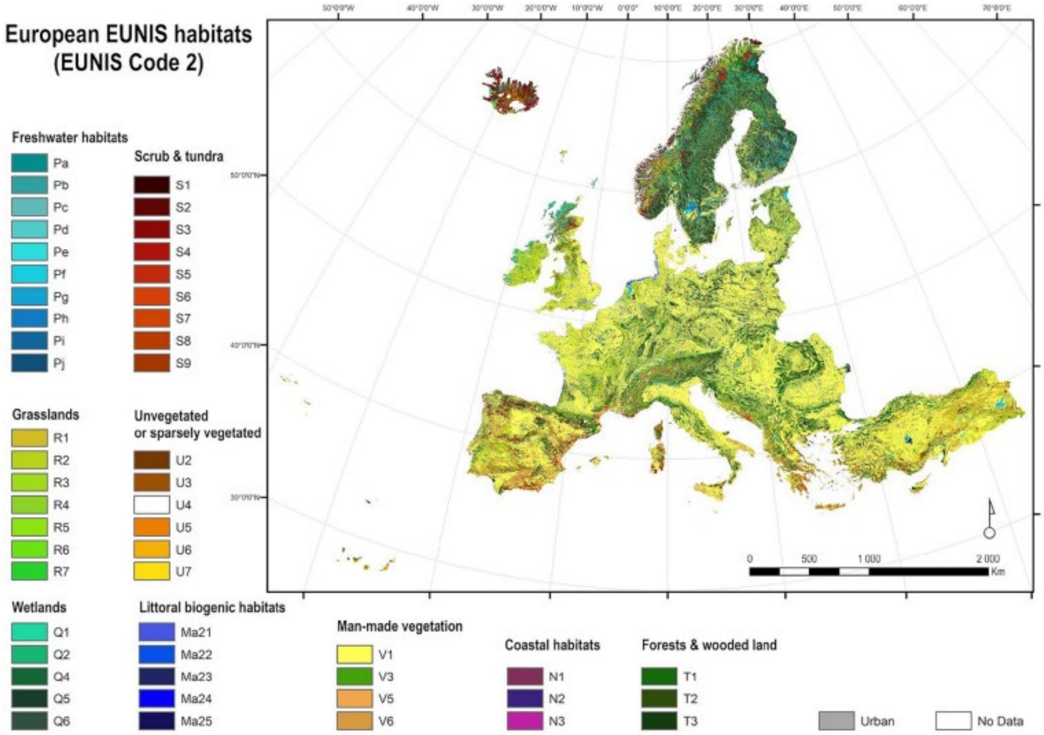
Figure 3.3 shows the main three step of the "evolution" from individual vegetation relevés to the final probability map of a EUNIS class: (1) distribution map, where georeferenced relevés are assigned to EUNIS classes using expert rules; (2) suitability map, where distribution data are related to climate, soil, and environmental data; and (3) probability map, where Copernicus high resolution layers (HRLs) were used to refine the suitability map. While the suitability map can be considered as a potential distribution map, the probability map presents a more precise distribution of the habitat type, although the probability map still represents a modelled distribution and may overestimate the actual distribution.

Figure 3.3 Illustration of the three steps, from individual vegetation relevés to a final probability map, with the EUNIS class "G1.6a: Fagus woodland on non-acid soils"



This approach was further developed by a research team, which produced an updated European EUNIS Habitat Map (Si-Moussi *et al.*, 2025). The study by Si-Moussi *et al.* (2025) models the distribution of EUNIS habitats across Europe using an ensemble of machine learning algorithms. The resulting 'European Habitat Map' (a modelled distribution estimate) covers 260 EUNIS habitat types at level 3 and has a spatial resolution of 100 m. For reasons of visibility, **Figure 3.4** illustrates the spatial distribution of these EUNIS habitats aggregated to the higher hierarchical level 2.

Figure 3.4. Wall-to-wall EUNIS habitat map visualized on the level 2 codes (Si-Moussi et al., 2025)



What makes this initiative particularly valuable is its use of the European Vegetation Archive (EVA) as training data, enabling the characterisation of habitats at a continental scale in a harmonised way. This approach directly addresses the challenges of existing habitat suitability maps, which typically suffer from coarse resolution and infrequent updates.

Land Cover Classification in Greece's Ionian Islands

Researchers in Greece have applied various machine learning techniques for land cover classification in the Ionian Islands region. Their study found that Random Forest (RF) algorithms delivered the best classification performance when incorporating multiple dataset types, including thermal data and terrain-related indices.

This project demonstrated AI's ability to bridge the gap between expert semantics and numerical satellite data. The research incorporated CORINE Land Cover product categories while achieving higher accuracy through machine learning. The classification included artificial surfaces, agricultural areas, forest and semi-natural areas, wetlands, and water bodies—precisely addressing the challenging ecosystem classes mentioned in the query.
Source: https://lup.lub.lu.se/student-papers/record/9147315/file/9147318.pdf

3.4.3 Pan-European AI initiatives for ecosystem mapping

The Geo-Harmonizer Project for Continental Europe

The Geo-Harmonizer project represents one of the most ambitious applications of AI for environmental monitoring across Europe. This initiative aims to map land-cover dynamics over the past two decades (2000 to 2020) for Continental Europe, using spatiotemporal machine learning models to predict land cover changes.

The project's approach involves collecting high-quality training samples from different years and integrating them with harmonized raster layers that correlate with land cover patterns across the EU. This

benchmark dataset provides valuable resources for training and evaluating different space-time models for land cover mapping across the continent. The temporal dimension of this project directly addresses the needs for monitoring ecosystem changes over time, which is essential for conservation planning and land-use decision-making.
Source: https://opendatascience.eu/geo-harmonizer/benchmark-dataset-mapping-land-cover-europe/

AI Ecosystem Mapping Initiative
While focused more broadly on mapping AI capabilities rather than natural ecosystems, the AI Ecosystem Mapping Initiative is a collaborative effort between VISION, the European Commission, and nine AI Networks of Excellence. This initiative aims to provide a comprehensive overview of AI-related research, development, and innovation activities across Europe.

The mapping tool ensures that ecosystem mapping continues sustainably through a Joint Topic Group named 'Ecosystem Mapping & Information Repository' (EMIR) under the AI-Data-Robotics association. Though not directly related to environmental mapping, this infrastructure demonstrates Europe's commitment to developing AI capabilities that could ultimately benefit environmental applications.
Sources: https://www.agimus-project.eu/news-events/news/45-launch-of-ai-ecosystem-mapping-initiative-with-ai-networks-of-excellence
https://www.vision4ai.eu/ecosystem-mapping/

AI4EU Platform for Resource Sharing
The EU-funded AI4EU project is building the first European AI On-Demand Platform and Ecosystem to share resources, tools, knowledge, and algorithms between Member States. While not specifically focused on ecosystem mapping, this platform provides a foundation for AI applications in environmental contexts. The project implements pilots led by industrial partners to demonstrate the platform's capabilities, potentially including applications for environmental monitoring. By lowering barriers to innovation and boosting technology transfer, AI4EU could accelerate the adoption of AI techniques for ecosystem mapping across Europe.
Source: https://cordis.europa.eu/project/id/825619

AI foundational models
Image embedding represents a significant advancement in the field of visual artificial intelligence. This technique enables the extraction of continuous vector representations from images, allowing machines to interpret and analyse visual data with greater accuracy and efficiency. By transforming images into numerical feature vectors, image embedding serves as a foundational process for machine learning models to better understand and process visual content.

Embeddings can be thought of as highly abstract, semantic summaries of images expressed in mathematical terms. While images are fundamentally composed of pixels, their semantic meaning emerges from the patterns and values within those pixels. Embeddings encode these abstractions directly as numerical vectors, effectively capturing much of the information required for feature detection. As a result, retrieving and computing semantic information from images becomes significantly faster and more efficient, since a substantial portion of the necessary computation is inherently embedded within these vectors.

4 Recommended production strategy

Building directly on the critical evaluation of baseline approaches (Chapter 2) and best practices identified from related initiatives (Chapter 3), this chapter presents the project team's **tailor-made proposal** for a next-generation CLMS data product supporting robust ecosystem accounting and mapping. The proposed methodology is explicitly designed to fill the technical and operational gaps highlighted in earlier chapters, offering a forward-looking pathway that responds to the recently updated policy framework at EU level.

The evidence collected throughout preceding chapters demonstrates the necessity of a specialised, integrated approach—one that leverages the proven strengths of existing CLC production while systematically introducing innovative enhancements, including advanced Earth observation tools, artificial intelligence methods, and harmonised in-situ data foundations.

- **Section 4.1** details the recommended solution—a bespoke workflow for ecosystem extent mapping—explaining how it transforms and upgrades the baseline model into a more flexible, efficient, and policy-relevant product.
- **Section 4.2** outlines an alternative fallback option, which, while less optimal, could serve as a contingency under certain constraints or transitional scenarios.

Taken together, the methodologies presented in this chapter synthesise all prior findings into a comprehensive proposal. This solution is directly informed by operational benchmarking (Chapter 2) and cross-initiative insights (Chapter 3), ensuring that the recommended path is both evidence-based and properly aligned with user and stakeholder requirements.

4.1 Tailored production approach

4.1.1 Introduction

The analysis presented in the previous chapters points toward a specialised approach that builds upon traditional CLC production methodologies, while:
- incorporating substantial enhancements to address current limitations,
- and modifying specific elements of the existing method to improve efficiency.

This approach has been refined through a detailed analysis of various processing steps and leverages multiple data sources, including targeted and consolidated layers such as the most relevant CLMS European datasets and technologies supporting ecosystem accounting. The core principle of this tailor-made methodology is to employ a unified (and more centralised) production process that generates a dedicated, purpose-built dataset designed to support coherent, reliable, and timely ecosystem accounts and related mapping products. However, the traditional CLC-production principle of working in cooperation with countries is maintained.

The creation of this purpose-built dataset is underpinned by three main pillars:
1. **Utilising Regional and Local Data and Expertise:** Integrating local datasets and country expert knowledge to enhance accuracy and provide contextual understanding of ecosystem characteristics unique to different regions.
2. **Integrating CLMS Data as Ancillary Datasets:** Employing existing CLMS products as supplementary datasets to enrich the baseline information and address potential gaps in primary data sources.
3. **AI Application and Infrastructure as a Service:** Strategically applying machine learning models where they add value—particularly for bridging data gaps, mapping challenging ecosystem classes, and supporting change probability detection. Offering these tools through a centralised cloud-based processing infrastructure, offered as a service (IaaS).

The following sub-sections address:
a) Proposed improvements to the current CLC production approach
b) Optimisation of the production process

c) Implementation of automation opportunities
d) Practical considerations for implementation

4.1.2 Key improvements over current CLC methodology

To meet the broader requirements for ecosystem extent products, the approach strengthens several critical aspects:

1. **Accelerate Production Timeline**: streamlining the entire business and system processes—from organization, procurement, and coordination through to final product generation—reduces the product update cycle from six years to three. The goal is to complete production within a maximum of 18 months from the reference year, without compromising quality.

2. **Enhanced Information content:**
    - Refined spatial resolution:
        - A differentiated MMU approach will be implemented, with a maximum 1-hectare MMU for urban ecosystem types and a maximum 10-hectare MMU for other ecosystem types. This requires careful balancing to ensure that the different MMUs align with overall requirements while maintaining both temporal and thematic consistency across dataset classes and time-series versions.
        - Drawing on lessons from the current CLC methodology, the updated approach proposes aligning the status and change layers to the same spatial resolution.
    - The approach aims to comprehensively cover all level 1 and, as much as possible, level 2 ecosystem types, while also providing options to identify specific level 3 ecosystem types.

4.1.3 Workflow optimisation

To achieve a three-year production cycle without compromising quality, it is essential to comprehensively optimise all workflow elements. These optimisations encompass technical, infrastructural, and administrative domains to establish an efficient end-to-end process (see **Table 4.1**).

From a technical perspective, the workflow relies on improved data acquisition strategies that enable faster collection of cloud-free imagery from multiple satellite sources. Sentinel-2 serves as the primary data source, supplemented by Sentinel-1 and additional platforms such as Landsat or very high-resolution (VHR) imagery available through the rapid response desk, as needed. Enhanced processing techniques further optimise image stacking and analysis, significantly reducing computational time while preserving analytical integrity.

Infrastructure optimisation centres on the implementation of cloud-based solutions, whether centralised or distributed, using Infrastructure as a Service (IaaS) models. This approach supports interoperable data management, including storage, access, processing, and extraction, thereby creating a seamless technical environment.

Beyond technical considerations, simplified administrative procedures are vital—streamlining tendering and contracting processes to minimise non-technical delays. Efficient quality control and internal verification procedures need to ensure optimised data extraction and validation, supported by streamlined inspection protocols. Finally, improved photointerpretation processes further expedite human interpretation by harnessing change alert data generated from Earth Observation (EO) and AI-based methods.

Table 4.1. Process optimisation elements

Optimization area	Elements	Expected benefits
Data acquisition	Multi-satellite approach (S2, S1 + Landsat); Prioritization of cloud-free imagery	Faster data collection; Reduced temporal gaps
Processing techniques	Optimised image stacking; Enhanced methodologies	Reduced computational time; Maintained analytical integrity
Could infrastructure	IaaS implementation (centralized/distributed); Interoperable data systems	Seamless data operations; Scalable processing capacity
Procurement	Aim to accelerate tendering and streamline contracting	Minimised non-technical delay; Improved project timelines
Quality control	Optimised and standardised data validation; Simplified inspection protocols	Maintained quality standard; Reduced QC bottlenecks
Photointerpretation	EO/AI-based change alerts; Enhanced visualisation tools	Accelerated human interpretation; Focused expert attention

4.1.4 Strategic automation implementation

The tailor-made approach can be further strengthened by introducing automated data extraction from Earth Observation (EO) sources, utilising supervised machine learning techniques supported by in situ data and CLMS products. This automation can be applied in two main scenarios:

- **Case 1 – Efficiency Enhancement**:
  When the extent of ecosystem types matches that of CLC classes, automated approaches can be used to increase efficiency while maintaining—or even improving—the quality of the output.
- **Case 2 – Novel Extraction**:
  In instances where the extent of ecosystem types cannot be derived from CLC classes, automated extraction from EO data offers new possibilities. Here, machine learning methods can identify ecosystem extents that are not otherwise captured using traditional CLC-based mapping.

Case 2, in particular, provides a significant opportunity to develop products that extend beyond the capabilities of CLC, addressing broader user expectations and meeting the evolving information needs of EU environmental policies.

<u>Feasibility assessment</u>

The development of such a product should begin with identifying the key elements and components necessary for mapping those L2—or even L3—ecosystem types that cannot be derived from CLC data. The process should ultimately clarify which ecosystem types fall into the following categories:

1. **Mappable using EO data**:

The ecosystem type can be identified and delineated using Earth Observation (EO) data alone, either through automated methods or photointerpretation.

2. **Mappable using EO and in-situ data**:

The ecosystem type can be depicted and mapped using a combination of EO data and in-situ data that reflect specific ecological characteristics and conditions.

3. **Mappable with in-situ data**:
The ecosystem type can only be mapped using local ancillary data.

Automated process development framework
When developing an automated system, it is important to consider the following technical limitations and requirements:

- **Comprehensive assessment of automated method limitations**:
  Consider the constraints of algorithms, processing capacities, and methodological boundaries to realistically determine what can be achieved.
- **Evaluation of Sentinel-2 and Sentinel-1 data constraints**:
  Assess the information content of S2 and S1 data, including spectral band availability, temporal frequency, and susceptibility to atmospheric interference, as these factors can limit the usability of satellite imagery sources.
- **Improvement of spatial resolution**:
  Enhance spatial resolution (minimum mapping unit) to meet target specifications, while ensuring comparability and consistency across time.

Methodological innovation framework
To advance the framework's capabilities, methodological innovations are essential. This innovation structure is supported by two main pillars (**Table 4.2**):

The first pillar establishes a robust **knowledge integration framework** that systematically incorporates domain expertise and existing knowledge bases into automated processes. This involves developing standardised procedures for collecting, validating, and integrating regional expertise and ground-truth information, as well as enhancing comprehensive workflows to optimise algorithms and validate results.

The second pillar focuses on leveraging advanced, **state-of-the-art techniques**. This includes exploring AI systems capable of making informed decisions based on available data and established rules, as well as applying pattern recognition methods that can identify relationships within data without relying on predefined classifications. Furthermore, the implementation of systematic procedures for assessing the accuracy and reliability of outputs—with minimal human intervention—serves to further enhance the overall effectiveness of the framework.

Table 4.2. Innovation components

Knowledge integration	Advanced techniques
processing threads for collecting local expert knowledge	machine reasoning approaches
systematic approaches for incorporating in-situ data, clarifying the specific role of the data in the support	unsupervised machine learning methods
data extraction, algorithmic tuning, and verification processes	automated validation and verification protocols

The tailor-made approach could consist of the following main processes:
1. **Coordination of Satellite Image Acquisition and Processing**:
The EEA (or another entrusted entity) organises and coordinates satellite image acquisition and processing, considering the role of ESA in the operation and continuity of satellite services.
2. **Cloud-Based Infrastructure for Data Management**:
The CLMS at EEA (or another entrusted entity) establishes a cloud-based infrastructure to be commonly used for training, data storage and processing. This infrastructure is accessible to local experts in participating countries and includes user-friendly interfaces for data extraction and validation. Sufficient capacity is reserved for the substantial training expected.

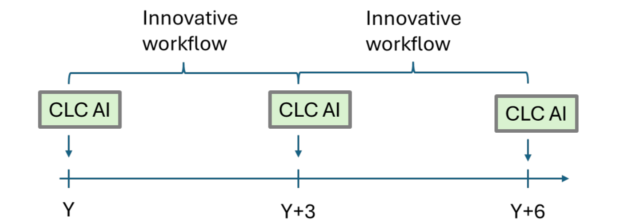
3. **Local Expert Contribution and In-Situ Data Integration**:
Country experts contribute local knowledge and expertise for data collection, photointerpretation, calibration, and validation. They also provide in-situ data to support the training of algorithms. Use is made of the legal and technical frameworks of the European Strategy for data to reduce the efforts in in-situ data collection and licensing.
4. **Collaborative Consultation and Feedback**:
Through the cloud-based infrastructure, the EEA (or another entrusted entity) engages with local experts to review and discuss the generated outputs. Experts provide feedback to enhance specific components of the processing chain, ideally enabling improvements to be incorporated automatically where feasible. Note: Machine reasoning can help minimise delays in human interaction and present new options to experts based on their feedback.
5. **Compilation and Production of the EU-Wide Product**:
The EEA (or another entrusted entity) compiles the validated outputs (mapped ecosystems per type and per country) and produces the EU-wide ecosystem extent product.

The involvement of local and regional experts is essential not only for ensuring the accuracy of data interpretation but also for capturing specific ecological, land use, and management practices that may not be detectable from EO data alone. Their expertise is critical for assessing the reliability of outputs, particularly when identifying nuanced or region-specific ecosystem types. This collaboration grounds automated and centralised processes in real-world, context-sensitive insight, thereby enhancing the robustness and acceptance of the final product. The centralised process is intended to relieve countries, as much as possible, from the mechanical aspects of the technical work.

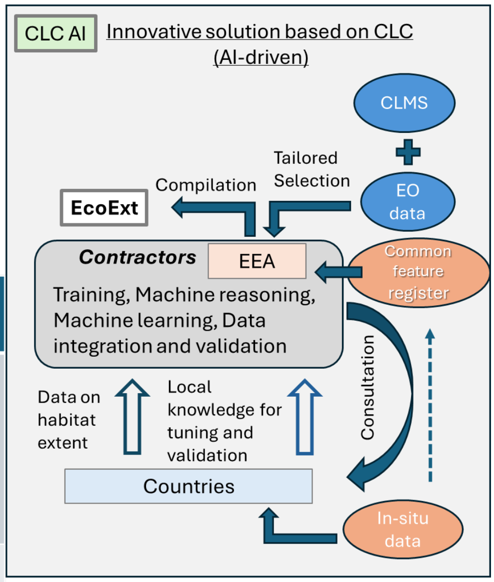
Figure 4.1 below illustrates the process, organised into three distinct sections: (a) Products generated by the approach along the reporting timeline; (b) Assessment and verdict of the approach; and (c) Schematic representation of the workflow for each individual product included in the approach.
The assessment and verdicts are based on the following criteria:

- **Fit for Purpose**: Provides the spatial foundation and thematic detail required for policy reporting and funding instruments (e.g., Ecosystem Accounting, NRR, Second Ecosystem Assessment).
- **Spatio-Temporal Consistency**: Ensures cross-country comparability, as well as accuracy and coherence across different time periods.
- **Endorsement by Countries**: Accepted and endorsed by countries as a potential basis for their reporting obligations.
- **Return on Investment (ROI)**: Evaluates the benefits realized within the required timeframe relative to the investment cost, also considering the complexity of operational deployment.

Figure 4.1. Tailored CLC-based approach. Top left: Products generated by the approach along reporting timeline; bottom left: Assessment and verdict of the approach; right: Schematic representation of the workflow of each individual product included in the approach



Approaches	Fit for purpose	Spatio-Temporal coherence	Interaction with countries	Cost
Assessment	Excellent, as CLC provides required thematic detail for all policies	High since same workflow is applied for each reporting year	Depends on the acceptance of the solution and the involvement of countries	Major investment required. Efficiency of AI approach is not yet proven
Verdict	🙂	🙂	😐	☹️



The tailored approach presented (Figure) offers the potential to deliver the most comprehensive information on ecosystems and related habitats; however, it requires high cost and investment. This necessitates the exploration of alternative solutions that are simpler, more practical, and efficient.

4.1.5 Practical considerations

The proposed workflow, drawing on successful elements of the CLC methodology, implements a structured, stepwise process in which inputs are harmonised and shared from source to final output. Each stage builds upon the previous one, with centralised management ensuring consistency, comparability, traceability, and continuity throughout the process.

Approach 1 represents the strategically optimal path forward: a comprehensive, tailor-made solution specifically designed to meet ecosystem mapping requirements. While this approach demands more substantial resources, it promises superior results, even though most of the proposed solutions will require practical testing and proof of concept. This section outlines key aspects that warrant deeper evaluation, ideally supported by practical exercises (proof of concept):

- It is essential to establish adequate and reliable basic spatial units that accurately reflect ecosystem assets. The suitability of the 2018 CLCplus Backbone vector product and OSM should be assessed for this purpose.
- The CLCplus Backbone time-series, produced biannually, could support the initial segmentation of primary land cover types, where appropriate, allowing for targeted and higher-frequency change detection.
- Similarly, HRL products (such as VLCC and NVLCC) can serve as complementary data sources to support CAPI, having demonstrated reliability in classifying key land cover types and related land uses (e.g., forest type, tree cover density, cropland, and artificial surfaces).
- Adherence to the principle of mutual exclusivity and collective exhaustiveness is critical for accurate class labelling and data topology; however, unstructured underlying data or information (such as the SIOSE approach used in Spain) is also valuable. It is important to retain access to this underlying information for evaluation purposes.

- Remapping the entire EU territory every reference year is unnecessary. Instead, an initial investment should be made in establishing accurate spatial units, followed by implementing effective monitoring to detect and account for changes, as advocated by the **'Change mapping first' strategy** described in Annex I.
- Further evaluation is needed to determine compatible MMU values, leveraging previous experience. A differentiated MMU approach is proposed: a maximum of 1 hectare for urban ecosystem types and 10 hectares for other types. Careful calibration is required to ensure alignment with requirements while maintaining temporal and thematic consistency across dataset classes and time-series versions.
- Economies of scale should be considered in budgeting for EU-wide deployment. The collection, compilation, and standardisation of in-situ data will likely represent the highest cost. It is worth exploring how AI can support and optimise the gathering of in-situ data (e.g., intensifying sampling in areas of potential change and reducing it in stable areas). The possible need for VHR satellite data from third-party missions should also be considered, with potential financial implications depending on the CLMS quota and the specific arrangements with ESA.
- For any AI-based solution, ongoing interaction with countries is essential. Two-way communication is mutually beneficial and builds on the strong collaborative legacy established through CLC activities.
- To achieve a real economy of scale from the adoption of automated tools in a centralised processing infrastructure and ensure consistency in mapping outputs across countries, the collected in-situ data (local, EU-driven) should be organised in a **common feature register.** It could be initially based on CLC geometries.
- To maximise return on investment, it is recommended to integrate experiences and coordinate plans with innovative AI solutions in the field of ecosystem extent mapping, both nationally and internationally (e.g., Geo-Harmonizer, PEOPLE-EA/WEED), to reduce duplication of effort and investment. The evaluation should also consider advancing to greater thematic detail (up to level 3 and beyond), and, where possible, aligning with other policy needs (e.g., the Habitats Directive).
- Although "priority area monitoring" products are geographically limited, they provide valuable thematic insights. For instance, the coastal zones product supports comprehensive mapping of coastal ecosystems and offers a certain level of reliability, provided it meets the required temporal references.
- Several cases described in Section 4 further highlight the value of CLMS products. In particular, CLC and the Coastal Zones products serve as foundational resources for AI model training or validation in the creation of new ecosystem data products. Notable examples include the EU Grassland Watch, the European Wetland Map, and the Ulmas-Walsh Map for Ireland.
- While the CLMS portfolio offers significant support, coordinating harmonised production timelines and ensuring product comparability will require additional effort. It remains important to consider existing semi-automated approaches for CLC updates. Some countries—such as Finland and Iceland (others may be added)—have successfully adopted these methods, which help address the complexity of monitoring specific land cover and land use classes. Additional relevant cases can be collected.
- Local experts involved in the current CLC implementation workflow can also contribute supplementary information as needed to meet the specifications required by ecosystem typologies (e.g., for mapping forest plantations), even if not originally foreseen in the existing CLC workflow. Specific examples should be identified and confirmed.
- Ensuring temporal coherence in the resulting product remains a key requirement.

4.2 Fallback approach: CLMS integration as a backup scenario

4.2.1 Purpose and position within the production strategy

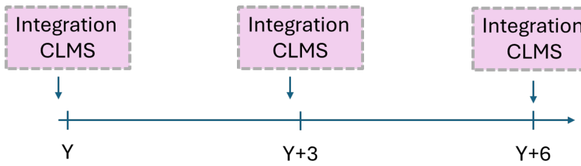
This section outlines an **alternative backup production strategy—not the primary or preferred approach—intended for use only if resource, timeline, or feasibility constraints prevent** implementation of the tailor-made, centrally managed solution presented in Section 4.1. The fallback approach should be considered a contingency: it aims to provide a minimum level of operational continuity and data availability, but comes with important caveats regarding quality, national endorsement, and long-term fitness for purpose.

This more conventional approach can leverage the integration of CLMS products with CLC to produce an ecosystem data product that fulfils user needs to the extent possible. As highlighted in previous chapters, for such an "integrated CLMS" approach to meet ecosystem extent product requirements, the following key elements must be addressed:

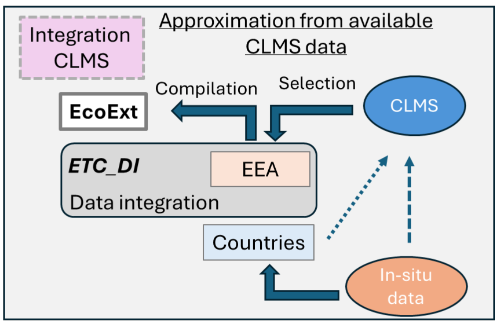
1. **Temporal Reference**: Ensure that all input CLMS products correspond to the same observation period required for the ecosystem extent product.
2. **Timeliness**: Guarantee that input CLMS products are produced sufficiently early—preferably within six months after the end of the reporting year—to enable the timely compilation of the final ecosystem extent product.
3. **Spatial Resolution**: Verify that the spatial resolution of input CLMS products meets the specified requirements.
4. **Thematic Content and Spatial Coverage**: Ensure that data integration and aggregation processes enable replication of CLC classes across the entire European territory, with consistency maintained among the various CLMS products.

Figure 4.2 below illustrates this process.

Figure 4.2. CLMS 'integration' approach. Top left: Products generated by the approach along reporting timeline; bottom left: Assessment and verdict of the approach; right: Schematic representation of the workflow of each individual product included in the approach



Approaches	Fit for purpose	Spatio-Temporal coherence	Interaction with countries	Cost
Assessment	Poor as the thematic detail is limited	Depends on the compatibility of the individual product instances	Low acceptance since countries are not involved	No new investment in data generation; only in workflows & data integration
Verdict				



Production is top-down and relies of the existing CLMS products

4.2.2 Summary of the fallback (CLMS Integration) approach

- Based on Section 2.1.2 and Chapter 2's evaluation, this backup strategy leverages the integration of existing Copernicus Land Monitoring Service (CLMS) data products and the Corine Land Cover (CLC) accounting layers, without substantial further enhancement or bespoke development. It represents an incremental improvement to existing methodologies rather than a tailor-made solution.
- The workflow entails combining ready-to-use CLMS layers (e.g., HRLs, Priority Area Monitoring datasets, CLCplus, Urban Atlas), aligning them as much as feasible to the EU ecosystem extent typology, and making necessary adjustments for aggregation and harmonisation.
- The result is a functional but sub-optimal outcome for ecosystem mapping needs, with some limitations for the upcoming accounting periods (e.g. some spatial areas will be represented by more classes than others, some classes will have a lower temporal accuracy than others), and no growth path towards a greater thematic detail.

4.2.3 Rationale for fallback use only

- This approach **does not meet all mandatory requirements** for the ecosystem extent product, as evidenced in Table 10 and discussed in Section 2.2.7. Key deficits include thematic alignment, completeness, and robust comparability.
- Because many CLMS layers were designed for purposes other than strict ecosystem accounting, their integration can only partially fulfil policy reporting needs.
- Most critically, the fallback approach carries an increased risk of limited acceptance and endorsement by Member States, which could undermine its utility and practical return on investment.

4.2.4 Enhancements ("Low-hanging fruit")—conditional on CLMS progress

As referenced in **Table** and Section 2.2, several incremental improvements could be made within this fallback framework to **mitigate certain known limitations**:

- Harmonise nomenclature and legends across CLMS layers with the ecosystem typology (see Section 2.2.2).
- Synchronise reference years and update cycles for constituent CLMS datasets to allow coherent change detection (see Section 2.2.3 and Annex I).
- Working towards the availability of CLMS data within 6 months after the reporting year, supporting timely national and EU-level accounting (referenced in Section 4.1.5).
- Expand spatial coverage and fill gaps in Priority Area Monitoring products where feasible (see Section 2.2.5).
- Unify or standardise the minimum mapping unit (MMU) across all CLMS inputs, aiming for the specifications used in ecosystem accounting (Sections 2.2.3, 4.1.2).
- Provide transparent, harmonised metadata and processing documentation for all integrated layers (see Section 2.2.4).

4.2.5 Limitations and additional Member State involvement

Despite the above improvements, core limitations remain that cannot be resolved within this integration-based scenario:

- Thematic and spatial gaps for certain ecosystem classes (see Section 2.2.2, Table).
- Incomplete consistency and comparability across time and between regions (Section 2.2.4, Table).
- Lack of a robust, systematic national quality assurance and endorsement process (Section 3.1 and 4.1.5).
- Reduced control over continuity due to diversified dependencies on evolving external datasets (Section 2.2.6).

Consequently, Member States would need to provide "partial-CLC" overlays or additional in-situ/national datasets for those classes or areas where CLMS products prove inadequate, and to engage actively in quality assurance and expert review.

4.2.6 Operational Implications and Strategic Conclusion

- **This fallback scenario is explicitly offered as a backup solution only.** It may offer a more resource-efficient or rapid interim product, but its success is constrained by technical and political limitations.
- Experience from previous evaluations (see **Table 2.8**) indicates that a solution without strong Member State buy-in, endorsement, and validation could ultimately be less valuable—potentially yielding a lower real-world return on investment than a more ambitious, centralised approach.
- The fallback approach serves as a stopgap: it can guarantee operational continuity should immediate deployment be required, but is not positioned as a path toward a fully policy-ready, fit-for-purpose EU ecosystem data product.

5 Conclusions and way forward

5.1 Aims of this study

This report is the first comprehensive analysis dedicated to the question of how an EU-wide ecosystem data product could be systematically developed and produced to meet emerging policy needs. Previous efforts have focused on isolated technical or thematic aspects, but this assessment now integrates policy requirements, current technical capabilities, and operational best practices into a single, coherent proposal. As such, it aims to provide a reference point for both immediate implementation and longer-term strategy in European ecosystem accounting and mapping.

5.2 Lessons learned from analysis of options

The evaluation of options has drawn on extensive technical, methodological, and policy expertise from both a pan-European and national perspective. Through iterative assessments and comparison with other projects, the study presents insights into:

- The true scope and complexity of producing a harmonised, fit-for-purpose ecosystem extent product.
- The strengths and lingering limitations of existing approaches, including both the CLC-based workflow and CLMS data integration.
- The indispensable role of in-situ data, human expertise, and advanced technologies (such as AI) within any scalable solution.

- The operational risks and production dependency factors that could impact the reliability, comparability, and continuity of these products moving forward.

5.3 Evaluation factors

The methods used in evaluating methodological choices considered the following core factors:

- **Thematic content and nomenclature alignment**: Ability to match the EU ecosystem typology at required levels, addressing policy and reporting demands.
- **Spatial resolution and minimum mapping unit (MMU)**: Capability to achieve fine-scale mapping in line with regulatory thresholds for both urban and non-urban systems.
- **Temporal frequency and consistency**: Alignment with update intervals and the need for robust, repeatable monitoring to detect real environmental change.
- **Completeness and MECE principle**: Adherence to the mutually exclusive and collectively exhaustive requirements for ecosystem asset delineation.
- **Continuity and operational dependencies**: Robustness of the approach in the face of evolving satellite missions, technology platforms, and organisational arrangements.
- **Validation, transparency, and endorsement**: Mechanisms for rigorous, independent validation; full documentation of methods; and acceptance at both EU and Member State level.
- **Integration of AI and automation**: Potential to increase efficiency and detail without compromising core validation and policy alignment requirements.

5.4 Recommendations

1. Adopt a tailor-made, centralised production model

- Build on the well-established CLC production framework, enhancing it for greater spatial resolution, thematic detail, and faster update cycles.
- Leverage the proven benefits of central coordination, mandatory standards, and national expert engagement to ensure pan-European consistency and user trust.
- Ensure the deployment of the production chain on cloud-based processing infrastructure offered as a service (CDSE, Green Deal Dataspace).

2. Systematically integrate CLMS, ancillary, and in-situ data

- Use CLMS datasets and emerging products as supplementary inputs, not as the primary foundation, to fill spatial or thematic gaps.
- Formalise and consolidate the collection and use of in-situ data and ancillary datasets through a common feature register for robust model training, validation, and verification.

3. Advance strategic use of AI and semi-automation

- Deploy machine learning/reasoning and automation in areas where it provides clear qualitative or efficiency gains—specifically for complex or underrepresented ecosystem classes.
- Develop protocols for AI transparency, validation, and integration of local expert knowledge.

4. Strengthen validation and iterative improvement mechanisms

- Create a separate, systematically resourced validation workflow, supported by harmonised in-situ and ancillary data collection.
- Establish iterative feedback loops with Member States and stakeholders to profit from local knowledge and enable rapid iteration and improvement based on operational experience.

5. Emphasise open documentation, governance, and capacity building

- Ensure all processes and methods are transparently documented and accessible, with clear linkages to INSPIRE and reporting requirements.

- Invest in training, technical infrastructure, and knowledge exchange to enable Member States to actively contribute to, and benefit from, the common mapping system.

5.5 Next Steps and EU-Level Decisions

5.5.1 Immediate Steps

- Pilot the tailored production workflow with selected Member States to verify technical assumptions, refine technical protocols and administrative processes.
- Fast-track technical working groups to define operational minimum mapping units, validation protocols, and AI integration points.
- Initiate a coordinated, EU-level in-situ data collection campaign aligned to the proposed workflow.

5.5.2 Next steps at EU-Level

- Ongoing EU-level dialogues on strategic directions for the Copernicus Land Monitoring Service (CLMS) need to review options for securing stable funding for a CLMS product that effectively supports ecosystem mapping in support of the EU ecosystem accounting module, NRR implementation and EU-level assessments.
- This review needs to consider funding for a centralised, iterative production model, ensuring the European Environment Agency (EEA) or another mandated entity has the mandate and resources to oversee the process.
- It will also be important to ensure the long-term availability of cloud-based processing infrastructure offered as a service (IaaS).
- Decisions need to be taken on the adoption of an upgraded and tailored CLC-based workflow as the reference method for ecosystem extent mapping at the EU scale, with binding minimum requirements for frequency, resolution, and thematic detail.
- An essential complement would be the development and resourcing of a pan-European in-situ data strategy, including standardised protocols, a common feature register and a protected database to support both AI model training and independent validation.
- Mandate the alignment of future CLMS developments and national mapping efforts with these new standards, enabling seamless integration for both EU-wide and national reporting obligations. Create an opportunity for countries to develop more detailed and local-specific ecosystem-related products and Copernicus downstream services.
- Facilitate stakeholder engagement, particularly with national statistical offices and ministries, to ensure political buy-in, operational feasibility, and user acceptance.

5.5.3 Meeting policy requirements and creating value-added

Delivering a robust EU ecosystem data product will require strategic deployment of resources and ongoing collaboration across organisational boundaries. Only when appropriate investment decisions are taken can ecosystem mapping data be delivered to meet EU policy needs, in support of the EU ecosystem accounting module, NRR implementation and EU-level ecosystem assessments. Furthermore, the traditional Corine land cover (CLC) data set is used as a central geo-spatial data platform in developing a multitude of other analytical products, whether in support of EU agri-environment indicators, land use modelling or other geo-spatial analysis. Ensuring continuity of this data series in a modernised form would guarantee a continuation of the long-term CLC time series and provide a central geo-spatial data platform with diverse functionality for many other applications and EU indicators.

Annex I: Consistency and comparability of EU-wide land cover/land use products – methods and limitations

Reliable analysis of natural processes requires consistent monitoring of individual land cover and land use elements. However, one of the greatest challenges in land monitoring is to establish comparability between LCLU products, both in terms of thematic comparability and temporal consistency.

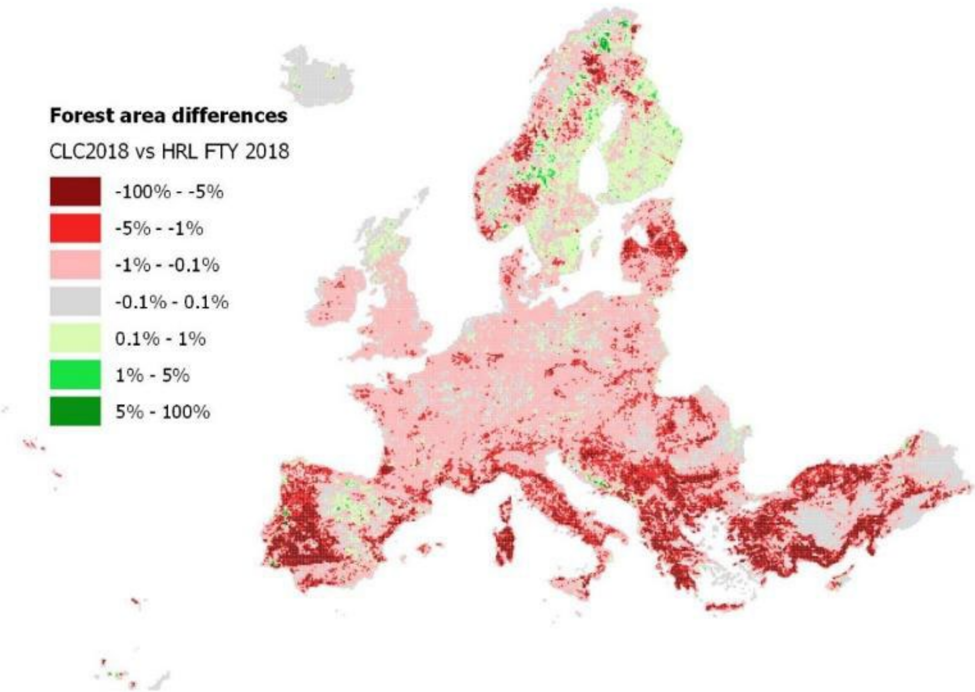
Thematic comparability

In general, there are small to large differences between the results of individual surveys, both in terms of features mapped and statistics derived. As a consequence, direct comparability of LCLU data is limited due to differences in:

- Semantic content (class definitions are often a mix of LC / LU / other characteristics).
- Surveying methods and instructions (visual interpretation, image classification, field survey).
- Scale (MMU, MMW, raster resolution).
- Quality and availability of input and reference data.
- Accuracy (criteria and realisation).

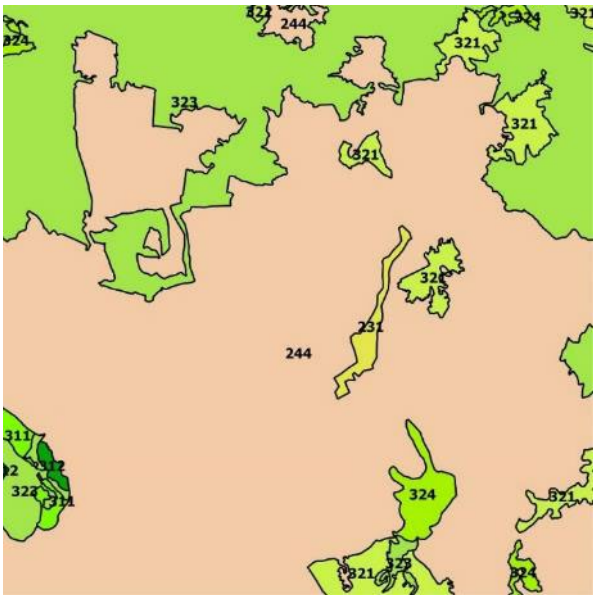
Thematic comparability issues are illustrated by the example of forest area mapped by Corine Land Cover vs high resolution raster data like HRL Forest Type or CLCplus Backbone (**Figure AI.1**).

Figure AI.1. Aggregated difference map of forest cover differences. Green colours indicate surplus of forest cover mapped by CLC2018 forest classes (311, 312, 313), while red colours indicate surplus of forest cover mapped by HRL FTY2018 (classes 1&2)

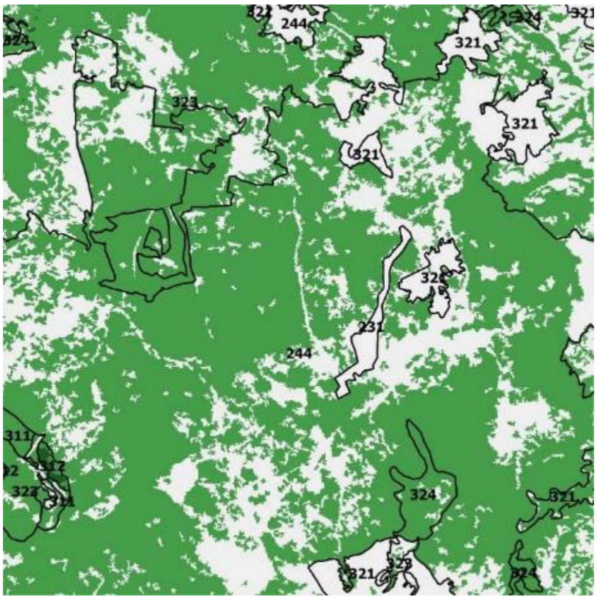


Forest areas mapped by various CLMS data were aggregated to 10x10km statistical grid. A slight surplus of mapped forest is shown by HRL FTY2018 for most of the continental part of Europe, while an extreme surplus of FTY mapped forest cover is indicated for the most part of the Mediterranean area, but partly for Baltic countries, Scandinavia, and other mountainous areas in Europe as well.

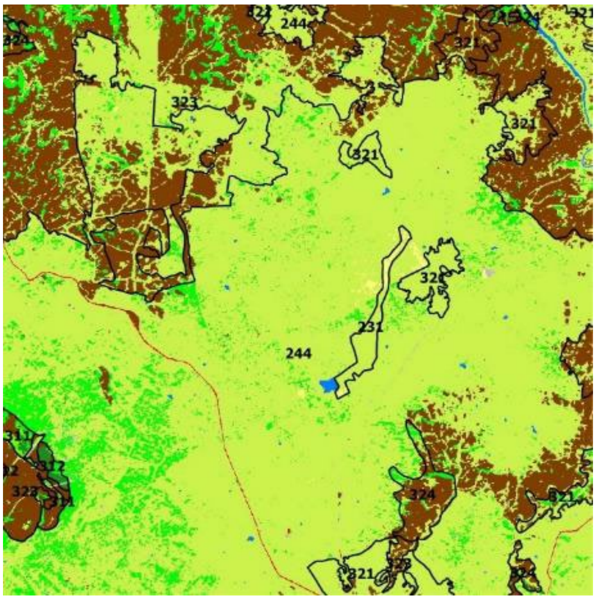
Figure AI.2. Forest area mapped by various CLMS surveys on agro forestry x,y = 2860482,1991771 close to Membrio, Spain



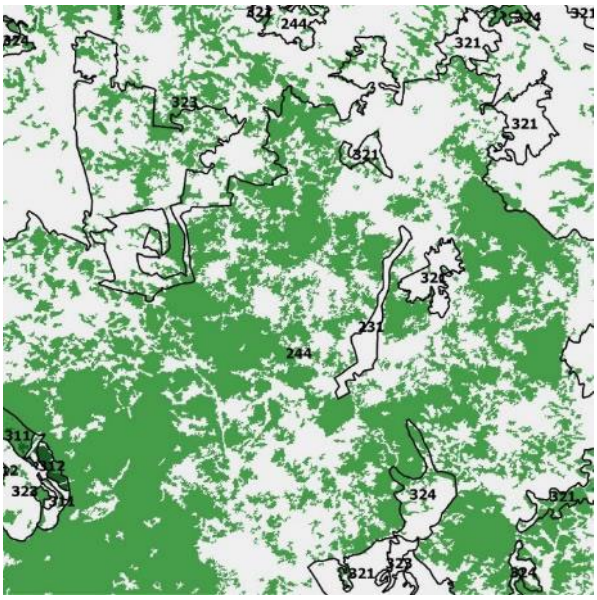
CLC2018 raster and outlines



HRL FTY2018 & CLC2018 outlines



CLCplus Backbone 2018 & CLC2018 outlines

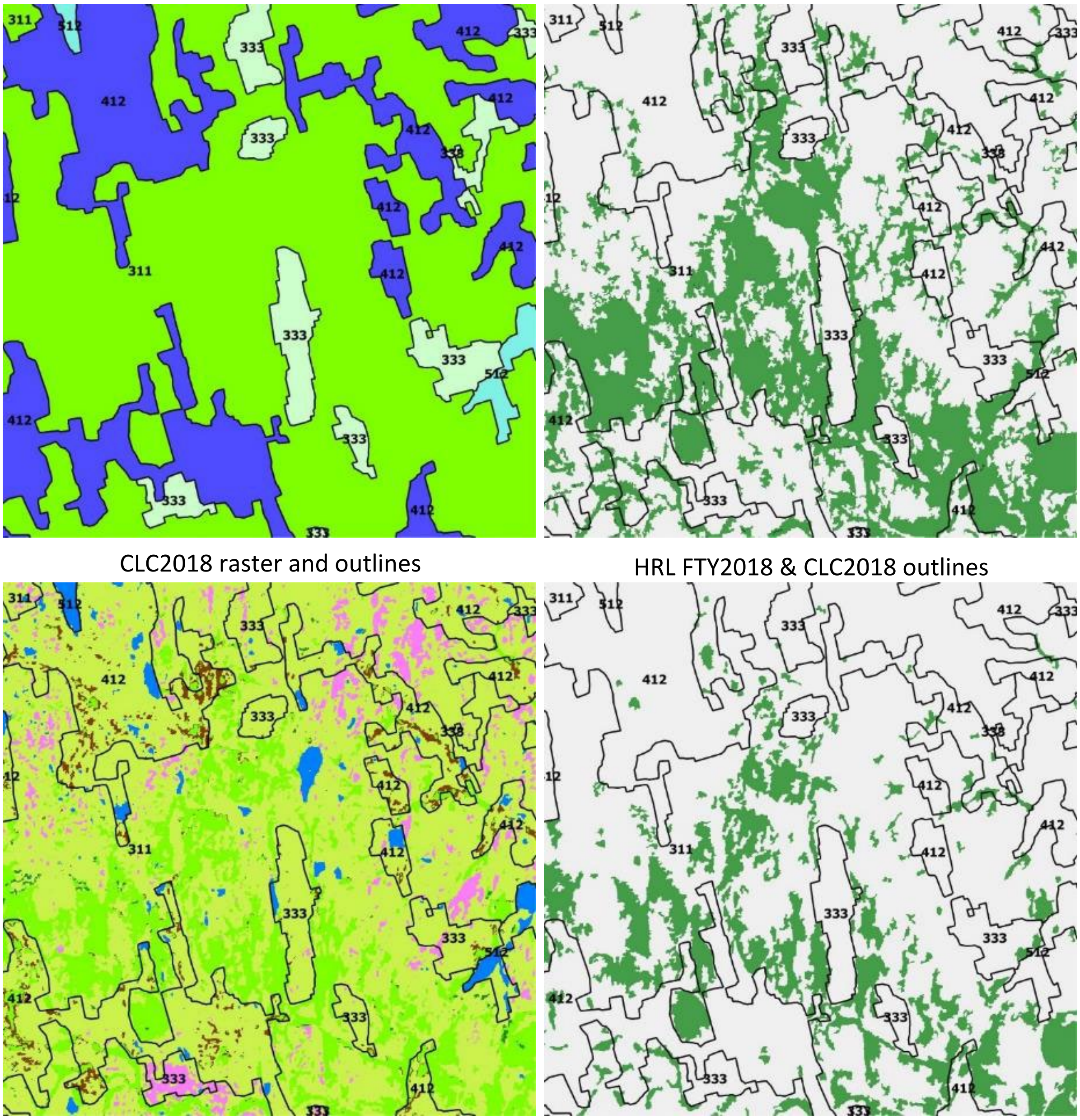


VLCC FTY2018 & CLC2018 outlines

Forest areas characterised by low density tree cover in Mediterranean areas are often mapped differently by various surveys:

- The land use aspect is dominating in case of CLC2018 by mapping large agro-forestry area (class 244) in Spanish example (**Figure AI.2**);
- FTY2018 layer created by HRL Forest 2018 survey mapped most part of CLC agro-forestry area as FTY forest by generalising low-density tree cover to larger contiguous forest blocks;
- The new VLCC Forest 2018 survey provides a revised FTY2018 layer. This layer indicates significantly less forest cover within CLC agro-forestry area;
- CLCplus Backbone 2018 survey indicates permanent herbaceous cover (class 6) for the most part of the CLC agroforestry area with sporadic broadleaved deciduous (class 3) patches. Sclerophyllous vegetation (CLC class 323) is mapped mostly as low growing woody vegetation (class 5) by CLCplus Backbone, mostly as forest by HRL FTY and mostly as non-forest by VLCC FTY.

Figure AI.3 Forest area mapped by various CLMS surveys on agro forestry x,y= 4898857,5177100 close to Soussjarvi, Norway



CLC2018 raster and outlines

HRL FTY2018 & CLC2018 outlines

CLCplus Backbone 2018 & CLC2018 outlines

VLCC FTY2018 & CLC2018 outlines

The aggregated difference indicates a surplus of forest cover mapped by CLC2018 for isolated locations, like in northern Scandinavia (**Figure AI.3**):

- Low density woody cover is mapped similarly by CLCplus Backbone and VLCC FTY layers by capturing area of relatively higher density of tree cover as woody / forest area;
- FTY2018 created by HRL Forest 2018 survey indicates significantly higher forest cover on the area;
- CLC2018 is indication broadleaved forest (class 311) for the most part of the area by generalizing smaller low density tree cover patches to larger forest area.

The examples above illustrate among other effects of scale differences (CLC vs FTY & Backbone), semantic differences (land use aspect by CLC vs land cover aspect by FTY & Backbone) as well as mapping concept and accuracy issues between FTY maps created by different surveys.

Temporal consistency

Consistency across time is one of the greatest challenges of EO-based products on land cover and land use monitoring. Temporal comparability is critical for accurate change detection and trend analysis.

Technical consistency

We say that a time-series is consistent in a technical sense, if the difference between status layers of any two reference years corresponds to the corresponding change layer:

Change = New status – Old status

Unfortunately, this simple requirement is violated in most currently available EU-wide LCLU time series because of several practical reasons, representing various levels of consistency, like:

- The difference of individual status is usually the sum of real LCLU changes and technical changes ("noise") due to classification uncertainties or subjectivity by visual interpretation;
- Application of different scales, minimum mapping unit or raster resolution for status and changes:
  - MMU for CLC status mapping is 25ha, while for CLC change mapping 5ha;
  - HRLs are mapped in 10m resolution since 2018, while change layers are still published in 20m resolution.

Real consistency

Real consistency would mean that if the time-series is consistent in a technical sense, and the presented changes represent real LCLU changes.

Concepts of change mapping

Two main strategies are applied currently to create CLMS change data:

I. <u>Change mapping first:</u> Vector-based LCLU products (CLC, Urban Atlas, Natura 2000, Riparian Zones, Coastal Zones);

II. <u>Post classification comparison:</u> Raster-based data (HRL Imperviousness, Forest, Grassland and SWF data).

The post classification comparison method delivers faster results, but the difference of status maps includes usually many technical (false) changes beyond real land cover changes, and the separation of these real changes from "noise" requires additional efforts.

<u>'Change mapping first' strategy</u>

This method was introduced in CLC2000 project by some of the countries and applied later by all of the EEA39 countries from CLC2006 inventory and upcoming updates. The method is based on the direct visual photointerpretation of change features based on Earth Observation (mainly satellite) imagery.

Main steps are:

1. Revision of the available status (correction in case of interpretation mistakes found);
2. Delineation and coding of meaningful change features considering pre-defined MMU and MMW;
3. Creating status for other reference year (either new or old status in case of backdating) by GIS operation combining available revised status layer and change layer.

The method was introduced by CLC mapping and applied by Priority Area Mapping products (Urban Atlas, Natura 2000, Riparian Zones, Coastal Zones) as well, however, with some variations.

'Post classification comparison' strategy

The post classification comparison change mapping method is applied for most of CLMS High Resolution Layers (HRL), supplemented by the semi-automatic separation of technical changes from real land cover changes.

Main and simplified steps of the complex processing are:

1. Creating status layer for first reference year by complex classification procedure;
2. Creating status layer for second reference year by complex classification procedure;
3. Creating difference layer by GIS procedure;
4. Separation of technical changes from real changes by semi-automatic procedures, creating change layer for publication.

This strategy was applied in several variations in the production history of HRLs, in some cases supplemented by additional harmonisation steps.

Time-series harmonisation

Consistency in a technical sense may be reached by applying harmonisation on the time-series as a post-processing.

Example of CLC accounting layers

A typical example is the time series of CLC accounting layers, where the latest (and presumably the best quality) CLC status is "back-dated" by combining this with CLC changes between previous periods to gain CLC status of previous reference years.

Harmonisation led to an increased comparability in CLC time series statistics, as many effects causing apparent false change signals are filtered out. The resulting data layers are known as accounting data layers and are used presently by EEA's Land and Ecosystem Accounting system[1].

The review of CLC harmonisation led to a revised methodology, which will be applied to create the new CLC accounting time-series after the completion of CLC2024 mapping.

Harmonisation of HRLs

Time-series harmonisation was applied in several variations in the production history of HRLs e.g.:

- Technical harmonisation of HRL Imperviousness time-series for the period 2006-2015. This harmonisation was carried out by Service Providers in the framework of the HRL 2015 survey, currently, the published layers are the harmonised ones;
- Technical harmonisation of the annual VLCC HRL Forest time-series for the period 2018-2021. This harmonisation was carried out by Service Providers in the framework of the VLCC 2021 survey, and included the revision of HRL Forest 2018 data;
- Harmonisation of soil sealing time-series for the period 2006-2018. The work performed in the frame of ETC/DI SC59425 "Support to the development of indicators and assessment of impacts of land use on the landscape scale" upon request from DG REGIO. Results include a harmonized time-series of soil sealing created on the basis of combining CLCplus Backbone status with Imperviousness Classified Change (IMCC) data[2];
- Harmonisation of HRL Forest time-series for the period 2012-2018. This more experimental kind of activity was performed with limited resources in frames of two different contracts:
  - Under AP2021 Task 1.2.4.2 "Forest assessment, supported by Copernicus and contributing to FISE";
  - In the frame of SC59425 "Support to the development of indicators and assessment of impacts of land use on the landscape scale" Task5: "Urban green infrastructure".

Example of CLCplus Backbone

The portfolio of CLCplus Backbone layers currently includes three status (2018, 2021 and 2023), but no change layers are published. The evolution of CLCplus Backbone status is currently being analysed by ETC/DI. Preliminary conclusions are as follows:

- CLCplus Backbone represent a high-quality land cover time-series with obvious signs of internal harmonisation effort;
- CLCplus Backbone is not fully harmonised in a technical sense; the differences include numerous 1-2 pixels sized transitions, most of these may be considered as noise due to uncertainty of individual classifications;
- The credibility of differences as a real change show various result depending on thematic classes and type of transitions appear.

Annex II: AI enhancement of the CLC approach – opportunities and integration pathways

Artificial intelligence (AI) presents a transformative opportunity for ecosystem mapping in Europe, offering greater accuracy, efficiency, and the ability to integrate vast amounts of environmental data. However, while AI can enhance mapping capabilities, its implementation is complex and requires addressing challenges related to data quality, model adaptability, and ethical considerations. A key area where AI can bring significant value is in complementing the Corine Land Cover (CLC) approach, helping to overcome its limitations in detecting certain ecosystem extent classes and refining spatial and thematic accuracy.

One of AI's greatest strengths lies in its ability to analyse and interpret remote sensing data, making it an ideal tool to enhance the CLC framework. Through machine learning for predictive modelling, AI can refine ecosystem classifications by filling in gaps where CLC struggles—such as detecting mixed agricultural landscapes, distinguishing broadleaved evergreen forests from deciduous ones, or differentiating artificial and natural water bodies. AI can also process satellite imagery at higher resolutions, improving the classification of small-scale landscape features such as greenhouses, cemeteries, and artificial shorelines that often fall below the CLC mapping threshold.

Another area where AI can support the CLC approach is AI-powered image analysis, which enables the automated detection and classification of vegetation, land cover changes, and ecosystem structures. Traditional CLC mapping relies on manual interpretation and standardized classification rules, which can sometimes lead to the misidentification of dynamic or complex ecosystems. AI, on the other hand, can learn from large datasets and adapt to variations, improving the detection of heterogeneous land cover types such as mixed agricultural patterns, tundra, and sparsely vegetated areas. Additionally, AI models can process aerial and hyperspectral imagery to distinguish between artificial and natural rivers, lakes, and coastal environments, addressing one of the key gaps in CLC-based classification.

Beyond image analysis, AI also excels in integrating multiple data sources, providing a more holistic approach to ecosystem mapping. By combining remote sensing data, climate variables, soil properties, and in-situ biodiversity observations, AI can help refine CLC-based classifications, particularly in cases where spectral similarities in satellite imagery make differentiation challenging. This is particularly valuable for coastal and wetland ecosystems, where AI can incorporate LiDAR elevation models, tidal data, and soil moisture indices to improve the delineation of land-sea boundaries and seasonal wetland dynamics.

Despite these advantages, AI is not a standalone solution—it requires human expertise and validation to ensure the accuracy and reliability of its outputs. Ground-truthing remains essential, as ecological complexity cannot always be captured purely through automated processes. AI-driven models should therefore be seen as a way to augment and refine the CLC-based process rather than replace it. By using AI to pre-process and enhance classification layers, the CLC workflow could become more efficient, consistent, and adaptable to evolving ecosystem changes.

Implementing an AI-enhanced approach within the CLC framework would require investment in computing infrastructure, data acquisition, and algorithm development. However, the potential benefits are substantial. AI could provide more accurate and standardized ecosystem maps, improving conservation planning and land-use decision-making. It could also enhance efficiency, processing vast amounts of data much faster than traditional methods. The long-term cost-effectiveness of AI could reduce the need for frequent manual updates.

In conclusion, AI offers a powerful set of tools to complement and enhance the CLC approach, addressing its limitations and making ecosystem mapping more precise and dynamic. A phased implementation approach, starting with pilot projects to integrate AI-based refinements into CLC workflows, could allow for iterative improvements while managing costs. By balancing AI's analytical power with human ecological expertise, Europe can achieve a more robust and adaptable ecosystem monitoring framework, ultimately supporting better environmental management and policy decisions.

Annex III: Wetland nomenclature challenges – scope, harmonisation, and data gaps

The map covers EEA38+UK, and it is based on a compilation and harmonisation of several global, European, and national datasets at 1 arcsec resolution (approx. 30m), distributed both as polygon and raster layers by country. This exercise, however, comes at the cost of a significant generalisation of the nomenclature, reducing the number of wetland classes to general levels, with a total of 10 classes corresponding to **4 wetland groups**, which differ considerably in terms of the characteristics and reference dates of the input data, being the following:

- **Floodplains:** this group includes a flood risk class with area that would be flooded during a 100-year flood, according to Dottori et al (2021)[12], the potential floodplain and the actual riparian zone, the last two based on the delineation made by the Copernicus Riparian Zones product.

- **Coastal wetlands:** it includes the classes tidal flat, estuaries, lagoon, salt marsh and salines according to the Copernicus Coastal Zones products, with additional data extracted from the Global Wetland Map[13] (only for the tidal flat, salines, and salt marsh classes)

- **Peatlands:** the peatland class builds on the update of the European Peatland Map from Tanneberger et al. (2017)[14] using data from over 100 different national and regional datasets.

- **Other wetlands:** this is a general 'wetland' class, containing wetlands that could not be clearly categorised in the other groups, most likely indicating mineral wetlands. Main source is the EEA Extended Wetland Ecosystem (EWE) layer 2012, including the following classes: *8 Managed or grazed wet meadows or pastures*, *9 Natural seasonally or permanently wet grassland*, *10 Wet heaths*, *11 Riverine and fen scrub*, *13 Inland marshes* and *14 Open mires*. Further wetland data was extracted from OpenStreetMap, such as reed beds, wet meadows, inland marshes, etc.

These 10 classes are the ones included in the raster dataset. The vector data are distributed in a geodatabase that includes additional fields on the wetland and peatland types. The main limitation is the lack of harmonisation of these classes, which are similar to those in the input data. This sometimes results in very basic categories or names in the national language. Despite this, the EMW is still a promising initiative, especially because of the great effort in collecting and reviewing data from such diverse sources. It should be noted that the scope of this product focuses on distinguishing between organic and mineral wetlands, as well as those that are managed or unmanaged in the case of peatlands. In fact, the map excludes water bodies and only keeps those that are specifically defined as wetlands by the input data. However, from an ecosystem accounting and mapping point of view, this results in a wetland map not as comprehensive as intended and may even be confusing or misleading at first glance, given the look and feel of the dataset.

On the one hand, floodplain classes are included like any other land cover type and have a very significant presence on the map, replacing information found in other existing datasets with the three floodplain classes, which are based on topographic or model-based delimitations. For example, river classes present in CLC, EWE layer or Copernicus Priority Area Monitoring (PAM) products are completely replaced by potential floodplain or actual riparian zone classes. In some cases, large areas of coastal lagoons or salt marshes fall into this category, which is not entirely incorrect but is confusing and much less representative of the actual wetland. The inclusion of these floodplain categories is interesting as they are potential wetlands, either because they are not currently mapped or because they have been transformed into

[12] http://data.europa.eu/89h/1d128b6c-a4ee-4858-9e34-6210707f3c81
[13] https://essd.copernicus.org/articles/15/265/2023/
[14] https://www.mires-and-peat.net/api/v1/articles/128692-the-peatland-map-of-europe.pdf

other types of cover (e.g. cropland or urban areas), but from the point of view of EEA and Copernicus mapping efforts, this is complementary information that should be distributed as ancillary data. Even in this sense, methods such as the one developed by Guelmami (2023)[15] within the Mediterranean Wetland Observatory[16] could be more interesting because it considers the potential wetland area throughout the whole territory, without being exclusively linked to floodplains, integrating data on topography, soils, surface water dynamics and current LULC cover.

On the other hand, the simplification of the nomenclature for those classes derived from Copernicus data and the EEA EWE layer (coastal wetlands and other wetlands) worsens the information currently available at the thematic level. The combination of high and low spatial resolution datasets also provides a confusing look and feel because some areas are more detailed than others. Furthermore, it is uncertain to what extent the additional data from the Global Wetland Map are relevant since it would require a more exhaustive validation process, but it is expected that there will be conflicts with the LULC products currently available, especially considering the global scope of this dataset.

The broader definition of peatland, considering all types of wetlands with organic soils, in addition to the inclusion of OpenStreetMap data, also generates quite a bit of confusion compared to previous work in the European context. In the reviewed areas, inland and salt marshes and rice fields are classified within this category. The peatland type defined in the geodatabase by OSM data for these areas is mostly incorrect or a very vague class. Although it is much more accurate at the spatial level, at the thematic level, it is inconsistent and less harmonised than what is available in the Copernicus and EEA products. It is also worth mentioning that, as a compendium of multiple data sources, the EWM contains information from several reference dates, resulting in some areas that may be more up to date than others. This is especially relevant in the case of OSM data, where it is difficult to identify the source of the data and track changes over time.

Figure AIII.1 and **Figure AIII.2** show a comparison between the EWE layer and the EWM to illustrate the different approaches to mapping wetlands and the inconsistencies that arise between the two datasets. The red circle focuses on L'étang de Berre, a coastal lagoon that is omitted from the EWM because it is incorrectly mapped as a lake in the Copernicus Coastal Zones product. It is therefore excluded from this map (since water bodies are not included) and classified as a potential floodplain. This is not false, but rather misleading and unrepresentative of the actual wetland coverage in the area. The yellow circle highlights a portion of the rice fields present in the Camargue region. Rice fields play a very important role in this region in terms of biodiversity, as they are a hybrid between wetlands and crops; hence, they are often considered wetlands from an ecological point of view. The different EWM approach omits this surface, classifying them with the three different floodplain classes. Again, this is not false, as these are frequently flooded areas, but they are not representative of the actual LULC. In other regions, the EWM identifies rice fields as peatlands because they are incorrectly mapped as marshes in OSM data. This is observed, for example, in the Albuera Natural Park in Valencia, Spain. Finally, the black circle highlights inland marshes mapped simply as wetland in EMW. Additional geodatabase information is required to find the "reedbed" category defined by OSM. **Figure AIII.1** shows this detailed classification according to the input data used to produce the dataset. EWE layer classes can be distinguished in the legend (e.g. Managed or grazed wet meadow or pasture, Open mires, Wet heaths). In this case, there is no harmonisation of the nomenclature, and classes are included as in the raw data.

[15] https://doi.org/10.1007/s41207-023-00443-6
[16] https://tourduvalat.org/en/newsletter-articles/la-cartographie-au-service-de-la-restauration-des-zones-humides-mediterraneennes/

Figure AIII.1. Detailed view of the Extended Wetland Ecosystem layer in the Camargue region, southern France

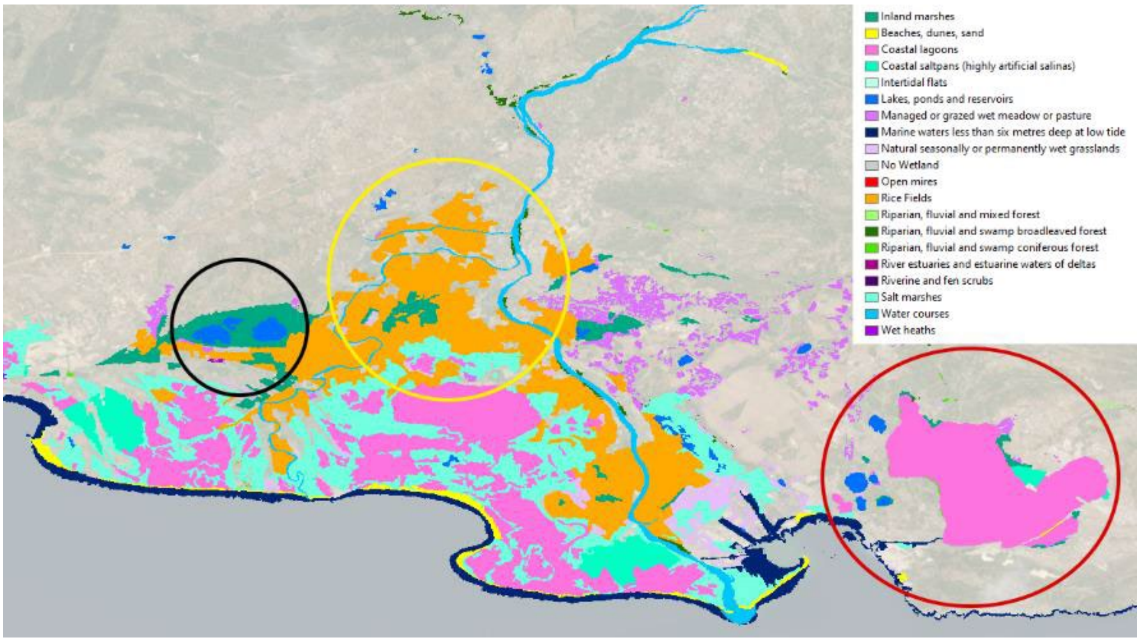


Figure AIII.2. Detailed view of the European Wetland Map in the Camargue region, southern France

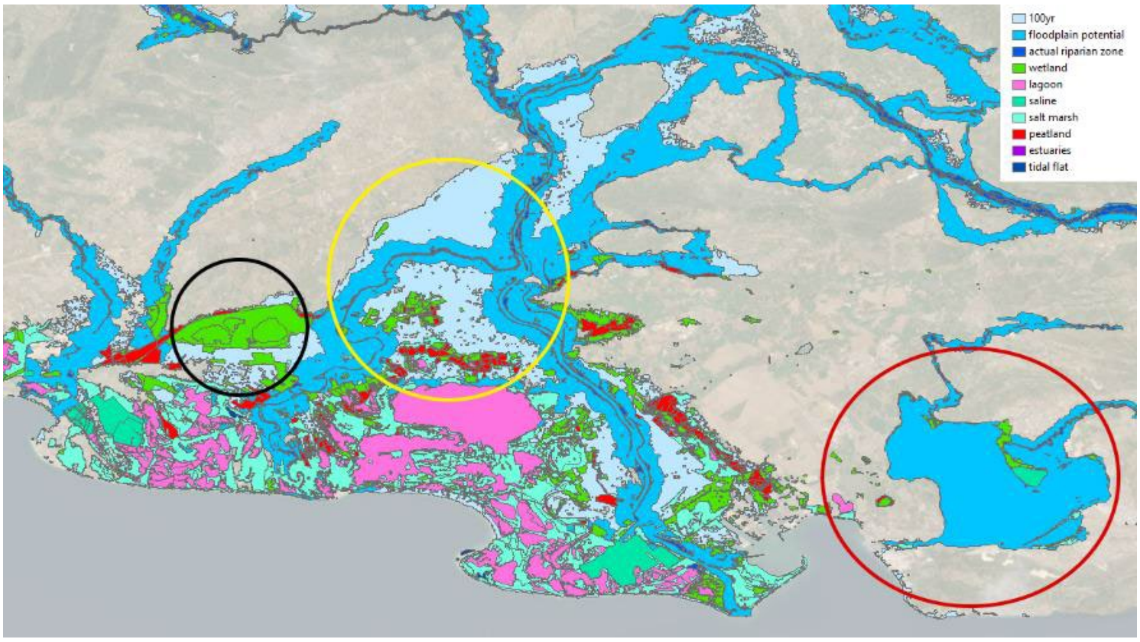
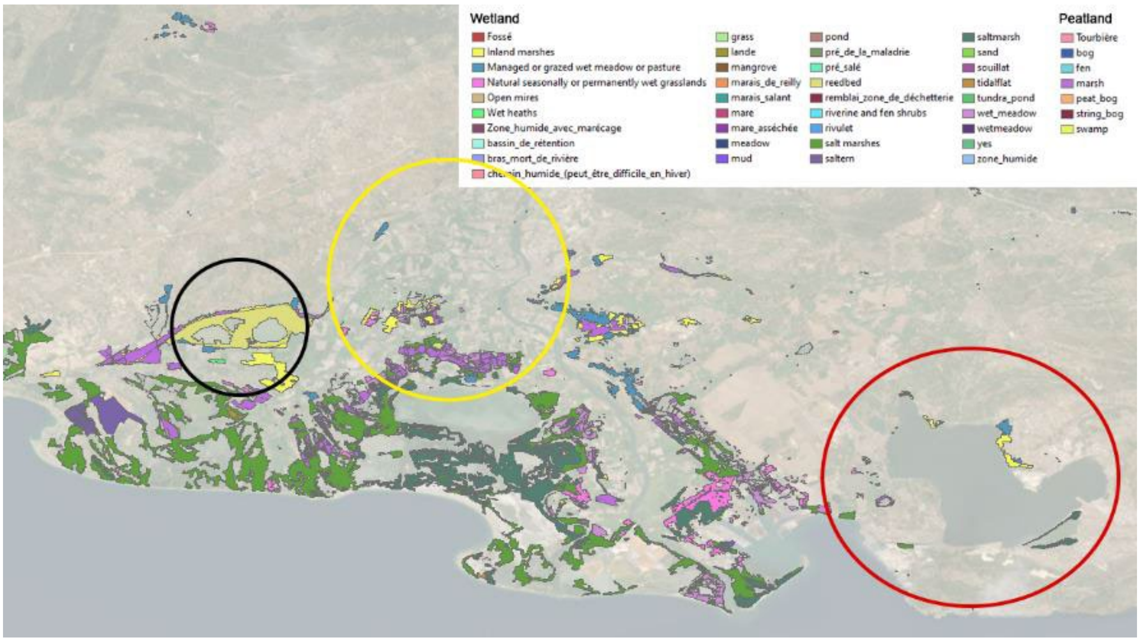


Figure AIII.3. Detailed view of the European Wetland Map in the Camargue region, southern France, showing wetland and peatland subclasses



In conclusion, the EWM is an interesting standalone mapping exercise, which gives a good overview of the state of the art in terms of GIS information available on wetlands at present date (**Figure AIII.3**). However, this initiative is linked to Horizon Europe projects, so resources are limited to their respective action plans, and the possibility of new versions very unlikely. In any case, as with many other mapping efforts, it relies on third-party input data, so the possibilities for updates and improvements are limited to future data availability scenarios. The sources of information for most wetland classes remain the current CLMS and EEA datasets, so the results and data updates present the same well-known problems and limitations.

Annex IV: Leveraging AI in European ecosystem mapping – potentials, requirements, and implementation barriers

Artificial Intelligence (AI) presents exciting opportunities for creating comprehensive and uniform ecosystem extent maps across Europe. However, this is not a straightforward endeavour, as it involves several aspects that require thorough discussion. This document examines the options, requirements, limitations, challenges, potential costs and benefits associated with using AI for this purpose.

1. Possibilities of AI in ecosystem mapping

AI provides several powerful approaches to create detailed ecosystem extent maps, namely:

1.1 Machine learning for predictive modelling

Machine learning algorithms can develop predictive models to estimate species distribution and habitat suitability. These models can:

- Utilise limited existing data to predict ecosystem extents in under-sampled areas
- Identify high-priority conservation zones
- Forecast climate change impacts on species and ecosystems

1.2 AI-Powered image analysis

AI excels at analysing various types of imagery to extract valuable ecosystem information:

- Process satellite and aerial imagery to identify and map specific habitat types
- Detect and classify vegetation, invasive species, and other ecosystem features
- Analyse underwater imagery for marine habitat mapping

This approach allows for rapid assessment of large areas that would be impractical to survey manually.

1.3 Integration of multiple data sources

AI can combine and analyse diverse data types to create more comprehensive ecosystem maps:

- Integrate remote sensing data, climate information, soil data, and other environmental variables
- Incorporate data from various sensors, including LiDAR, multispectral, and hyperspectral imagery
- Utilise environmental DNA (eDNA) data to detect species presence in ecosystems

1.4 Automated feature extraction

AI algorithms can automatically extract relevant features from various data sources:

- Identify and map forest cover and types from satellite imagery.
- Map soil properties and classifications using multispectral imagery.
- Identify and monitor wetlands and their seasonal variations.
- Derive vegetation indices and land cover classifications from satellite imagery
- Generate seascape derivatives from bathymetry data for marine ecosystem mapping

1.5 Continuous learning and updating

AI systems can continuously improve ecosystem maps as new data becomes available:

- Incorporate new field observations and remotely sensed data to refine existing maps
- Adapt to changing environmental conditions and ecosystem dynamics over time
- Identify areas where additional data collection is needed to improve map accuracy

2. Required input data

To train AI models for ecosystem mapping, various types of input data are necessary:

- **Remote sensing data**: This includes satellite imagery, aerial photographs, and LiDAR data. Satellite imagery provides broad coverage and can capture various spectral bands, useful for identifying different vegetation types and land cover. Aerial photographs offer high-resolution images for detailed analysis, while LiDAR data provides precise elevation and vegetation structure information.

- **Environmental variables**: These encompass climate data, soil information, and topography. Climate data includes temperature, precipitation, and other weather-related variables that influence ecosystem dynamics. Soil information covers soil types, moisture content, and nutrient levels, which are crucial for understanding plant growth and habitat suitability. Topography data, such as elevation and slope, helps in modelling water flow, erosion, and habitat distribution.
- **Species occurrence data**: This data comes from observations recorded during field surveys and citizen science projects. Field surveys provide accurate and detailed records of species presence and abundance, while citizen science projects can offer a large volume of data collected by volunteers, enhancing spatial and temporal coverage.
- **Existing ecosystem maps**: These maps are used for model training and validation. They provide a baseline of known ecosystem extents and types, helping to calibrate AI models and ensure they produce accurate and reliable predictions.
- **Ground-truth data**: Field observations are essential for validating AI model outputs. Ground-truthing involves comparing AI-generated maps with actual on-site observations to verify the accuracy of predicted ecosystem boundaries and features. This step ensures that the models are reliable and can be trusted for decision-making.

3. Limitations and human intervention

Despite AI's capabilities, human intervention remains crucial in several aspects:

3.1 Interpretation of complex ecological processes

Humans possess the expertise to interpret complex ecological interactions that AI may not fully understand. Ecologists can provide insights into subtle ecological processes, species behaviour, and the cultural significance of ecosystems.

3.2 Ground-truthing and validation

Field surveys and observations by experts are essential to validate AI predictions and models, ensuring accuracy and reliability. Ground-truthing involves verifying AI-generated maps with actual on-site observations to confirm ecosystem boundaries.

3.3 Ethical and cultural considerations

Humans bring an understanding of ethical, social, and cultural dimensions to ecosystem management that AI lacks. Conservation efforts often require consideration of local knowledge and cultural values, which AI alone cannot account for.

3.3 Adaptive management and decision-making

Human decision-makers use AI-generated insights to make informed choices about conservation strategies, land use planning, and resource management. Policymakers might use AI data to draft regulations that balance ecological preservation with economic development.

3.4 Addressing uncertainties and limitations

Experts are needed to address uncertainties and limitations in AI models, adapting them to specific ecological contexts. Climate change predictions often require expert interpretation to apply to local ecosystems accurately.

4. Challenges and complexity

Several challenges need to be addressed when using AI for ecosystem mapping:

4.1 Data quality and availability

AI models rely heavily on the quality and availability of data, which can be sparse or unevenly distributed in certain regions. Improving data collection methods and collaborating with local communities to gather comprehensive datasets is crucial.

4.2 Complexity of ecological interactions

Ecosystems are complex systems with numerous interacting variables, making it challenging for AI to capture all dynamics accurately. Employing hybrid models that combine AI with ecological theories can help better understand these interactions.

4.3 Transferability across regions

AI models developed in one region may not be directly applicable to other regions with different ecological conditions. Adapting models to local contexts by incorporating region-specific data and expert knowledge is necessary.

4.4 Ethical considerations

AI-driven decisions may overlook ethical considerations related to conservation and land management. Incorporating ethical frameworks and stakeholder perspectives in AI applications for ecosystem management is essential.

4.5 Interpretability and transparency

AI models can be complex and difficult to interpret, leading to challenges in understanding the rationale behind predictions. Developing interpretable models and transparent methodologies is crucial to enhance trust and acceptance.

4.6 Data integration and harmonisation

Combining diverse datasets from multiple countries and sources presents a significant challenge. This includes:

- Harmonising data formats and standards across different European countries and regions.
- Integrating various types of data (e.g., satellite imagery, field surveys, climate data)
- Ensuring consistent data quality and resolution across the continent

4.7 Model development and adaptation

Creating AI models that work effectively across Europe's diverse ecosystems is complex:

- Developing models that can accurately classify a wide range of habitat types
- Adapting models to account for regional variations in ecosystems
- Ensuring transferability of models across different geographical areas

4.8 Validation and ground-truthing

Verifying AI-generated maps across such a large and diverse area is challenging:

- Coordinating ground-truthing efforts across multiple countries
- Addressing discrepancies between AI predictions and local expert knowledge
- Continuously updating and refining models based on new data

5. Costs and benefits

Implementing an AI solution for Europe-wide ecosystem mapping would involve significant **costs**:

5.1 Infrastructure and computing resources

- High-performance computing systems to process vast amounts of data
- Storage solutions for large datasets and model outputs
- Cloud computing services for distributed processing and data access

Data Acquisition and Processing

- Purchasing or acquiring rights to high-resolution satellite imagery
- Costs associated with field surveys and data collection
- Data preprocessing and cleaning
-

5.2 Human Resources

- Salaries for AI specialists, data scientists, and ecologists
- Training programs for staff across different European countries
- Coordination and management of international teams
-

5.3 Ongoing Maintenance and Updates

- Regular model updates and refinements
- Continuous data collection and integration
- System maintenance and upgrades

Despite the complexity and costs, implementing an AI solution for ecosystem mapping across Europe offers significant **benefits**:

- **Improved Decision-Making**: More accurate and up-to-date ecosystem maps can inform better conservation and land-use policies.
- **Efficiency gains**: AI can process vast amounts of data much faster than traditional methods, potentially saving time and resources in the long run.
- **Standardisation**: A unified AI approach could help standardise ecosystem mapping across Europe, facilitating better cross-border collaboration and policymaking.
- **Early warning systems**: AI models could help detect ecosystem changes early, allowing for more proactive conservation efforts.
- **Cost-effectiveness in the long term**: While initial implementation costs may be high, the system could prove cost-effective over time by reducing the need for frequent manual surveys and analyses.

6. Conclusion

In conclusion, while AI offers powerful tools for creating comprehensive and homogeneous ecosystem extent maps for Europe, it is most effective when combined with human expertise. By addressing the challenges and limitations, we can leverage AI's strengths to enhance our understanding and management of European ecosystems.

Although implementing an AI solution for ecosystem mapping across Europe would be complex and costly, the potential benefits in terms of improved ecosystem management, conservation efforts, and policymaking could justify the investment. The key is to approach the implementation in phases, starting with pilot projects in specific regions and gradually scaling up to cover the entire continent. This phased approach would allow for iterative improvements and help manage costs and complexity over time.

Annex V: Habitat mapping vs. modelling – concepts, methodological distinctions, and practical implications

The term "habitat" has several meanings. In ecology, it means either the area and resources used by a particular species (the habitat of a species) or an assemblage of animals and plants together with their abiotic environment. This is the locality in which a plant or animal naturally grows or lives. In the spatial data context, habitats are geographical areas characterised by specific ecological conditions, processes, structure, and (life support) functions that physically support the organisms that live there. They include terrestrial, freshwater and marine areas distinguished by geographical, abiotic, and biotic features, whether entirely natural or semi-natural. A habitat is characterised by a relative uniformity of the physical environment and close interaction of all the biological species involved.

The INSPIRE directive for the establishment of spatial data infrastructure in Europe regards the spatial data on both habitats and biotopes as characterisation of the distribution of geographical areas being functional areas for living organisms: habitats being the spatial environment of specific species, while biotopes being the spatial environment of a biotic community (organisms living together in mutual dependence). Although conceptually different, biotopes and habitats are dealt with similarly when it comes to their spatial data capturing representation. A biotope can contain more than one habitat.

The general scope for the collection of data on habitats is nature conservation, mainly through the designation and management of protected sites to preserve endangered species and habitats. In addition, there are reporting obligations about the occurrence and abundance of habitats set in the different EU policies. This scope formulates the following key information needs:

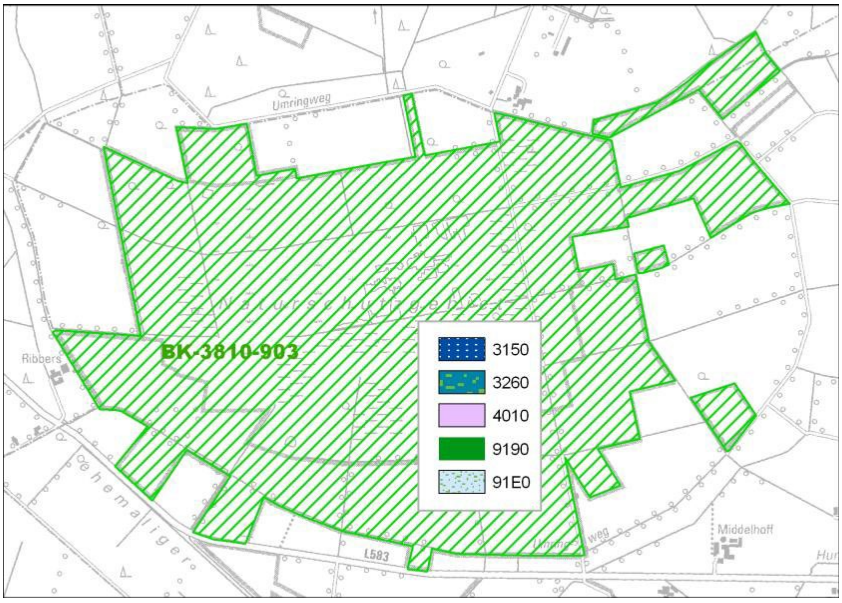
- geographic location and extent (area, length and/or volume) of habitats
- the geographic distribution of habitats

Habitats have natural boundaries and are classified according to their ecological or geo-physical conditions. Habitats are usually mapped based on fieldwork (most frequently) and/or remote sensing image interpretation (e.g. aerial photography interpretation) and sometimes also modelling. They have boundaries of their own rather than being defined relative to some other spatial object type.

Contrary, the boundaries of the distribution of the habitats and biotopes are not based on the habitat and biotope itself, but on the boundaries of other geographic features. Habitat distribution is usually depicted or modelled using as reference other spatial objects as analytical units, such as grid-cells (very frequently), bio-geographical regions, nature conservation sites or administrative units.
Therefore, the "map" of habitat and the "map" of habitat distribution should be considered two different geospatial products.

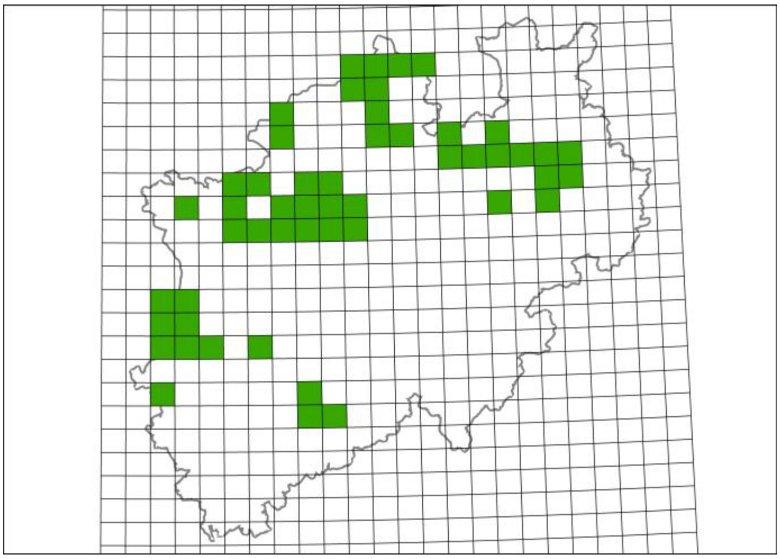
The habitat map is a georeferenced dataset holding a collection of non-overlapping discrete spatial objects (**Figure AV.1**). The spatial object (feature) of the habitat map is a geometry that depicts the natural boundary of a given habitat, or the extent of the area occupied by several habitats (habitat complex). The geometry of the spatial objects is usually a 2-dimensional polygon. Each spatial object holds several attributes recording the types of habitats present (according to a pre-defined nomenclature) and their area. It can hold, among other biophysical characteristics, information on the type of species and vegetation present.

Figure AV.1. Habitat map with one feature (habitat complex): this habitat feature contains 5 habitat types according to the pre-defined nomenclature; a list of these habitat types in the corresponding feature metadata will have further information on the total area



The habitat distribution map is a georeferenced dataset holding a complete tessellation of non-overlapping discrete spatial objects. The spatial object (feature) of the habitat distribution map is based on external reference features, which act as analytical units (**Figure AV.2**). These units could be a regular georeferenced grid or administrative units. Each such analytical unit holds information about the occurrence and abundance of a given habitat type, expressed as the number of individual habitats found and/or the area they cover within the given analytical unit.

Figure AV.2. Map of the distribution of habitat type 3130 referring to the analytical units of a Gauß-Krüger grid in North Rhine-Westphalia, Germany



Annex VI: The critical role of In-situ data in ecosystem extent mapping – foundations, protocols, and challenges

Key Messages and Principles

- **High-quality in-situ data**—with accurate spatial, temporal, and thematic attributes—are mandatory for the credible validation of ecosystem extent products and must be collected directly in the field.
- **A standardised protocol** for sampling and collection is essential. This should build on established EU protocols (e.g., for habitat ground truthing), integrating lessons from international methodologies such as IUCN GET identification keys.
- **Harmonised schemas** must be adopted by Member States, with validated in-situ points stored in a protected central database. This ensures they are reserved for validation and not accessed for model training, maintaining statistical independence and integrity.
- **Sampling schemes** should prioritize collection of 'change' reference points and address seasonal variability, ensuring that dynamic processes and temporal patterns are reflected.
- **Field collection tools** (e.g., mobile apps or standardised paper forms) should support consistent and robust reference data gathering and qualification.
- **Protocol alignment and integration** with existing data (LUCAS, Fluxnet, pollinator surveys) and anticipated mechanisms (such as EBOCC) is highly recommended to maximise efficiency and comparability.
- **Ensure semantic interoperability** through the adoption of standard meta-languages for machine-enables translation of classification systems, nomenclatures and vocabularies.
- The **protocol must facilitate EO-based (Earth Observation) algorithm requirements**, specifying spatial sample representativity, time of sampling, and appropriate metadata to enable direct integration with remote sensing workflows.
- For **training data**, a separate, lower-quality in-situ collection protocol is needed, allowing use of expert-controlled visual interpretation (e.g., VHR imagery, ecological datasets), citizen science (GBIF), and similar sources.
- **LIDAR campaigns** managed by countries represent a valuable auxiliary resource and should be systematically considered in the broader in-situ data infrastructure.

Both the capturing and mapping of ecosystem types—and the subsequent validation of the outcomes—aim to detect either a state or a change in the state of observable land surface phenomena. While the mapping process can increasingly rely on automated methods, **validation must depend on human judgment supported by field-based evidence**. In-situ data reflecting on-the-ground conditions are essential for both mapping and validation, as they help interpret how land phenomena are best observed and classified.

In particular, **in-situ data used for validation must be completely independent from data used in the training and tuning of Earth Observation (EO)-based models**, to preserve statistical independence and ensure unbiased accuracy assessments. In data-scarce situations, however, training datasets may indirectly support validation efforts—for example, by identifying similar areas for inspection—but must never be reused as reference points for validation (Congalton and Green, 2019).

In the context of ecosystem and habitat mapping, in-situ data refers to **direct observations of the current or changing state of land phenomena**, collected through structured procedures applied in their natural setting. These can take the form of individual points, mapped polygons, or transect-based observations. Desirable in-situ data characteristics include:

- **Systematic coverage**, ensuring spatial representativity and minimising sampling bias;
- **Affordability**, avoiding high acquisition and processing costs;

- **Timeliness**, with low latency between real-world changes and observation, and collection at ecologically appropriate moments.

While no field data collection method is flawless, **known limitations can be managed if metadata is comprehensive and transparent**. Metadata should describe both technical parameters (e.g. timing, equipment, resolution) and **epistemological context** (e.g. interpretation assumptions, classification thresholds). For instance, labelling an agricultural area as "abandoned" requires that the criteria behind that classification be clearly defined in a way that reflects local socio-environmental realities. This is especially crucial when observations convey interpreted information, not merely raw measurements.

Human observation remains one of the most intuitive ways to assess land conditions, whether in person or through proxies like panoramic imagery, drone-based surveys, or street-view technologies. These terrestrial methods, while often logistically demanding and time-sensitive, offer irreplaceable insights.

Computer-Assisted Photointerpretation (CAPI) offers a complementary approach, where trained interpreters use remote sensing imagery (e.g. VHR, HHR, Sentinel) and visual cues to derive ground-level insights. Despite lacking direct ground presence, **a skilled photo-interpreter can produce reference-quality data**, especially when supported by structured tools and protocols. The quality of such data depends on image characteristics (e.g. spatial and temporal resolution, spectral range) and coverage extent.

Conventional ground surveys and remote sensing–based methods are often carried out under highly standardised conditions, as with the LUCAS survey. However, **newer data sources from citizen science, volunteered geographic information (VGI), and geotagged imagery** offer increasingly valuable supplementary inputs. When appropriately filtered and processed—especially with AI tools like convolutional neural networks—these large, unstructured datasets can support ground-truthing and model development.

The modular design of the LUCAS survey (e.g., grassland, landscape features) provides a strong foundation for tailored ecosystem mapping approaches. Still, **some modules require advanced training to avoid ambiguity**, such as in estimating percentages of INSPIRE Pure Land Cover Components (PLCC) within a given observation frame. Clear and unambiguous geolocation, along with contextual information like weather or light conditions at the time of observation, is also vital to ensure data reliability.

Table AVI.1 below illustrates certain aspects of the possible ground data collection methods in support of ecosystem mapping.

Table AVI.1. The advantages and disadvantages of the various ground data collection methods and related data sources

Source/method	Pros	cons
Field data collection	Intuitive and universal (applicable to all land types)	High cost, time- and resource-intensive
StreetView-technology	Low or no cost	Coverage limited to accessible areas
Sentinel 1+2 / L8 CAPI	Freely available	Moderate resolution
HHR CAPI	Better spatial resolution	Acquisition cost applies
VHR CAPI	Highly detailed, familiar to experts	High cost and timing constraints
Aerial CAPI	High detail	Expensive, timing limitations
Geotagged imagery	Abundant and diverse	Quality and validity vary
Third-party data (national thematic data, Copernicus PAM, LUCAS)	Existing, often harmonised	Use case dependent

In-situ data also play a **direct role in delineating habitat or ecosystem boundaries**, particularly in areas where EO-based interpretation is complex or ambiguous. Despite advances in AI and cloud-based processing, satellite data still generates immense volumes of information that require structured filtering. **Using EO signals tied to predefined spatial units (rather than raw pixels)** reduces noise and allows for clearer, more manageable analyses. This tiered and reductive approach facilitates object-based mapping of ecosystems through logical decomposition of the landscape.

To support EU-wide environmental monitoring, there is a **pressing need for a common geospatial reference framework**. Segmenting very high-resolution imagery into homogeneous land units—each with a unique identifier—would allow multiple EO products and in-situ datasets to be **linked at the feature level**, significantly improving consistency and interoperability across monitoring initiatives.

References

Chytrý M, Tichý L, Hennekens SM, et al. EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Appl Veg Sci*. 2020 ;23:648–675. (Eleni, Rania and I are among the co-authors, along with now retired ETC expert Douglas Evans). https://doi.org/10.1111/avsc.12519

Clasen, K. N., Hackel, L., Burgert, T., Sumbul, G., Demir, B., & Markl, V. (2024). reben: Refined bigearthnet dataset for remote sensing image analysis. arXiv preprint arXiv:2407.03653.

Congalton, R.G., and Green, K. (2019). Assessing the Accuracy of Remotely Sensed Data Principles and Practices, Third Edition. CRC Press. pp. 346. ISBN: 9780429052729. https://doi.org/10.1201/9780429052729

Dobrinić, D., Miler, M., & Medak, D. (2024). A comparative analysis of pixel-based and object-based approaches using multitemporal PlanetScope imagery for land cover classification. Proceedings of the International Conference of Environmental Remote Sensing and GIS, 175–178. https://doi.org/10.5281/zenodo.11643869

Duro, D. C., Franklin, S. E., & Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sensing of Environment, 118, 259-272. https://doi.org/10.1016/j.rse.2011.11.020

EEA briefing from 2017, Underpinning European policy on nature conservation - Revision of the EUNIS habitat classification, https://www.eea.europa.eu/en/analysis/publications/underpinning-european-policy-on-nature-conservation

European Commission, Checks by Monitoring Quality Assurance Framework, ANNEX IV, Ground data, version 1.2, June 2021, Based on JRC 119733

Eurostat, EU ecosystem typology, Technical Note, December 2024

Eurostat, Guidance note on ecosystem extent accounts, Technical Note, December 2024 https://ec.europa.eu/eurostat/documents/3859598/23563646/Guidance-note-ecosystem-extent-accounts-2024.pdf

Hydro International. (n.d.). Marine Habitat Maps from Hydrography and Machine Learning. Retrieved from https://www.hydro-international.com/content/article/marine-habitat-maps-from-hydrography-and-machine-learning

Investment in artificial intelligence in the National Recovery and Resilience Plans https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/762288/EPRS_BRI(2024)762288_EN.pdf

IPBES. (n.d.). ARIES for IPBES. Retrieved from https://www.ipbes.net/policy-support/aries

Kucharczyk, M., Hay, G. J., Ghaffarian, S., & Hugenholtz, C. H. (2019). Geographic Object-Based Image Analysis: A Primer and Future Directions. *Remote Sensing*, *12*(12), 2012. https://doi.org/10.3390/rs12122012

Lei, Z., Bingfang, W., Yu, L., ·Xiaosong, L., Quanzhi, Y. (2024). Object-based multi-scale recognition approach for residential landscape classification. Asian Conference on Remote Sensing (ACRS 2024).

Milenov, P., Sima, A., Lugato, E., Acquafresca, L., Anastasakis, K. and Devos W., Deliverable 3 – Prototype and technical guidance for EO-based monitoring of peatland/wetland, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/61663, JRC134131.

Milenov, P.; Sima, A.; Lugato, E.; Devos, W.; Loudjani, P. Enabling Spatial Data Interoperability through the Use of a Semantic MetaModel—The Peatland Example from the JRC SEPLA Project. Land 2024, 13, 473. https://doi.org/10.3390/ land13040473

Miro. (n.d.). Ecosystem Mapping AI Assistance Template. Retrieved from https://miro.com/miroverse/ecosystem-mapping-ai-assistance-template/

Mücher, S. and Hennekens, S., 2019. Processing European habitat probability maps for EUNIS Forest (T), Heathland, scrub and tundra (S) and Grassland (R) habitat types based on vegetation relevés, environmental data and Copernicus land cover. ETC/BD report to the EEA.

NOAA. (n.d.). Enhancing Habitat Mapping Accuracy and Efficiency Using Artificial Intelligence. Retrieved from https://coastalscience.noaa.gov/project/enhancing-habitat-mapping-accuracy-and-efficiency-using-artificial-intelligence/

PEOPLE-EA, Evolution Roadmap (Deliverable 19), Contract No. 4000139263/22/I-AG, V1.4, 2025

Powell, R., Colin N. B. (2008). Land Use Land Cover Mapping in the Tiffin River Watershed.

Rodwell, Evans and Schaminée, 2018, Phytosociological relationships in European Union policy-related habitat classifications, https://doi.org/10.1007/s12210-018-0690-y

Sayre, R., Karagulle, D., Frye, C., Boucher, T., Wolff, N. H., Breyer, S., ... & Possingham, H. (2020). An assessment of the representation of ecosystems in global protected areas using new maps of World Climate Regions and World Ecosystems. Global Ecology and Conservation, 21, e00860.

Si-Moussi, S., Hennekens, S., Mücher, S., De Keersmaecker, W., Chytrý, M., Agrillo, E., Attorre, F., Biurrun, I., Bonari, G., Čarni, A., Ćušterevska, R., Dziuba, T., Ecker, K., Güler, B., Jandt, U., Jiménez-Alfaro, B., Lenoir, J., Svenning, J. C., Swacha, G., & Thuiller, W. (2025). EUNIS habitat maps: enhancing thematic and spatial resolution for Europe through machine learning. *Scientific Data*, *12*(1), Article 1976. https://doi.org/10.1038/s41597-025-06235-7

Sima A. et al. Use of geotagged photographs in the frame of Common Agriculture Policy checks, Ispra, 2020, JRC120223

Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019, July). Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. In IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium (pp. 5901-5904). IEEE.

Sumbul, G., De Wall, A., Kreuziger, T., Marcelino, F., Costa, H., Benevides, P., ... & Markl, V. (2021). BigEarthNet-MM: A large-scale, multimodal, multilabel benchmark archive for remote sensing image classification and retrieval [software and data sets]. IEEE Geoscience and Remote Sensing Magazine, 9(3), 174-180.

Thomson, E. C. (n.d.). Artificial Intelligence in Ecological Mapping. LinkedIn. Retrieved from https://www.linkedin.com/pulse/artificial-intelligence-ecological-mapping-thomsonec

UNEP. 1997. World atlas of desertification. United Nations Environment Programme.

VITO Remote Sensing. (n.d.). Habitat Mapping with AI. Retrieved from https://blog.vito.be/remotesensing/habitat-mapping-ai

World Economic Forum. (2024). AI in Conservation: Where We Came From and Where We Are Heading. Retrieved from https://www.weforum.org/agenda/2024/03/ai-in-conservation-where-we-came-from-and-where-we-are-heading/

Ye, S., Pontius, R. G., & Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. ISPRS Journal of Photogrammetry and Remote Sensing, 141, 137-147. https://doi.org/10.1016/j.isprsjprs.2018.04.002

Glossary

- **Advanced Earth Observation Processing**: Techniques using satellite data to improve the accuracy and efficiency of ecosystem mapping.
- **AI-Based Ecosystem Mapping**: The application of artificial intelligence technologies to enhance land cover mapping and classification processes.
- **Basic Spatial Units (BSU)**: Geometric constructs used to represent individual ecosystem assets in ecosystem accounting, such as grid cells or polygons.
- **BigEarthNet**: A dataset or project related to Earth observation and ecosystem mapping.
- **CLC (Corine Land Cover)**: A traditional methodology for land cover mapping in Europe, often enhanced with AI-based improvements.
- **CLCplus**: A framework for harmonized and tailored land cover information of Europe — Copernicus Land Monitoring Service.
- **Collectively Exhaustive**: A principle ensuring that spatial data covers the entire territory without omission.
- **Completeness**: A requirement ensuring that all relevant data is included in ecosystem mapping and accounting.
- **Consistency and Comparability**: The ability to ensure statistical comparability over time between reporting periods.
- **Copernicus Land Monitoring Services (CLMS)**: Ready-to-use data products provided by the Copernicus program for land monitoring purposes.
- **Ecosystem Accounting:** A framework for organizing spatial and statistical data to measure ecosystem extent and changes over time, aligned with EU environmental accounts regulation (691/2011).
- **Ecosystem Extent**: The spatial area of an ecosystem asset projected onto a 2D plane, measured in hectares, representing contiguous non-overlapping units.
- **ETRS89/LAEA Europe**: The coordinate reference system widely used for pan-European spatial analysis, with limitations in local-scale applications due to distortions.
- **EU Ecosystem Typology**: A hierarchical classification system for ecosystems in Europe, with Level 1 being mandatory and Level 2 preferable for reporting purposes.
- **Frequency**: Reporting frequency for ecosystem accounts, typically every three years.
- **Grassland Watch**: An initiative focused on monitoring grassland ecosystems in Europe.
- **Kunming-Montreal Global Biodiversity Framework (KM-GBF)**: An international framework defining biodiversity indicators, including the extent of natural ecosystems.
- **MECE Principle (Mutually Exclusive and Collectively Exhaustive)**: A principle ensuring that each spatial unit is assigned to one ecosystem type without overlap or omission.
- **Minimum Mapping Unit (MMU)**: The specific size of the smallest feature that is being reliably mapped. The target MMU for ecosystem mapping is defined as 1 hectare for urban areas and 10 hectares for non-urban areas.
- **Nature Restoration Regulation (NRR)**: EU legislation aimed at restoring degraded ecosystems and improving biodiversity.
- **Object-Oriented Approach**: A methodology in Earth observation that focuses on identifying objects within satellite imagery for classification purposes.
- **Pan-European AI Initiatives**: Collaborative projects across Europe leveraging AI for ecosystem mapping and monitoring.
- **PEOPLE-EA/ARIES FOR SEEA**: Projects integrating AI technologies into ecosystem accounting frameworks like SEEA.
- **SEEA EA (System of Environmental-Economic Accounting – Ecosystem Accounting)**: A global framework for integrating ecosystem data into environmental-economic accounts.

- **Spatial Resolution**: The level of detail captured in spatial data, defined as 100x100 meters for ecosystem accounting purposes.
- **Temporal Consistency**: Ensuring statistical comparability over time between reporting periods (e.g., every three years).
- **Thematic Accuracy**: The accuracy of mapped thematic details, targeting a minimum of 85%.

List of abbreviations

AI	Artificial Intelligence
AP	Annual Programme
API	Application Programming Interface
BSU	Basic Spatial Unit
CAPI	Computer-Assisted Photo-Interpretation
CDSE	Copernicus Data Space Ecosystem
CIS	Copernicus In Situ Component Information System
CLC	CORINE Land Cover
CLCAC	CORINE Land Cover Accounting Layers
CLCC	CORINE Land Cover Change
CLMS	Copernicus Land Monitoring Service
CNN	Convolutional Neural Network
DEM	Digital Elevation Model
DG	Directorate-General
EAGLE	EIONET Action Group on Land monitoring in Europe
EBOCC	EU Biodiversity Observation Coordination Centre
EEA	European Environment Agency
EFG	Ecosystem Functional Group
EO	Earth Observation
ESA	European Space Agency
ET	Ecosystem Type
ETC	European Topic Centre
ETRS	European Terrestrial Reference System
EU	European Union
EUGW	EU Grassland Watch
EUNIS	European Nature Information System
EVA	European Vegetation Archive
EWE	Extended Wetland Ecosystem
EWM	European Wetland Map
FISE	Forest Information System for Europe
GBF	Kunming-Montreal Global Biodiversity Framework
GBIF	Global Biodiversity Information Facility
GET	Global Ecosystem Typology
GIO	GMES Initial Operations
GIS	Geographic Information System
GLC	Global Land Cover
HD	Habitats Directive
HHR	High-High Resolution
HR	High Resolution
HRL	High Resolution Layer

IMCC	Imperviousness Change Classified
INCA	Integrated Natural Capital Accounting
INSPIRE	INSPIRE — Infrastructure for Spatial Information in the European Community
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
ISRIC	International Soil Reference and Information Centre
IUCN	International Union for Conservation of Nature
JRC	Joint Research Centre
LAEA	Lambert Azimuthal Equal-Area
LAI	Leaf Area Index
LC	Land Cover
LCC	Land Cover Change
LCFM	Land Cover and Forest Monitoring
LCLU	Land Cover / Land Use
LLM	Large Language Model
LPIS	Land Parcel Identification System
LUCAS	Land Use/Cover Area frame Survey
LULC	Land Use / Land Cover
LULUC	Land Use and Land-Use Change
LWC	Leaf Water Content
MAES	Mapping and Assessment of Ecosystems and their Services
MECE	Mutually Exclusive, Collectively Exhaustive
ML	Machine Learning
MMU	Minimum Mapping Unit
MMW	Minimum Mapping Width
MS	Member State
NDVI	Normalized Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
NRR	Nature Restoration Regulation
NVLCC	Non-Vegetated Land Cover Characteristics
OBIA	Object-Based Image Analysis
OSM	OpenStreetMap
PAM	Priority Area Monitoring
ROI	Region of Interest
SCL	Scene Classification Layer
SEEA	System of Environmental-Economic Accounting
SVM	Support Vector Machine
SWF	Small Woody Features
UN	United Nations
UNEP	United Nations Environment Programme
UNFCC	United Nations Framework Convention on Climate Change (UNFCCC)
VGI	Volunteered Geographic Information
VHR	Very High Resolution

VLCC	Vegetated Land Cover Characteristics
VPP	Vegetation Phenology and Productivity
WEED	World Ecosystem Extent Dynamics
WMS	Web Map Service

European Topic Centre on
Data integration and digitalisation
https://www.eionet.europa.eu/etcs/etc-di

The European Topic Centre on Data integration and digitalisation (ETC DI) is a consortium of European institutes under contract of the European Environment Agency.

European Environment Agency
European Topic Centre
Data integration and digitalisation