The Romanian Interoperability case study for landscape features

Report under the
EXPERT CONTRACT CONTRACT NUMBER
- CT-EX2018D319529-106

January 30 2022.

Author: Bernadett Csonka

Involved Institutions: Romanian Agriculture Payments and Intervention Agency - APIA[2]

Contacts: csonkadetti@gmail.com

Contents

Foreword

This report presents an analysis how the mapping of Landscape Features could be extended to all potential types of LFs defined on the entire eligible agriculture area of the Member State to support a higher level of biodiversity services in the new CAP framework. The study focuses on the challenges of improving the LF layer for the new PMEF-based CAP, also with the possibility to integrate the Copernicus Small Woody Feature layer as an external datasets.

Acknowledgements

This study would not have been possible without the Romanian Paying Agency (APIA) made available all the IACS-GIS databases for the area of pilot sites, I would like to extend my sincere thanks to contributing colleagues, Traian Crainic and Marian Dumitrescu. Special thanks to Katalin Tóth and colleagues from GTCAP team of EC- Joint Research Centre (JRC) for their valuable assistance and to Gábor Gulyás and Péter Kovács for their technical support.

Abstract

The main objective of the project is to complete on pilot sites the mapping of all potential landscape features linked to agricultural area, according to a harmonized and simplified definition and typology and to compare the mapped features to the Copernicus SWF database.

As part of analysis, the semantic mapping between the proposed broad categories of LF, the existing GAEC/EFA-LF categorization and the corresponding feature types of the created LF layer was completed. Mapping rules and LF definitions were tested and evaluated with visual interpretation (CAPI) and the effective use of simplified definition and functional typology proposed by the recently published "Landscape features in the EU Member States" technical report was successfully proven. Proposal for less specific, comprehensive definition of landscape features regarding the use of average width and land cover dominancy has been elaborated, and successfully tested. Matching of the digitized LFs with the Copernicus High resolution Small Woody Feature data was analyzed by a confusion matrix and by derived accuracy measures. The SWF dataset was not found to be adequate to derive directly the woody LF elements from, but its use to determine the area % of woody vegetation inside the area of a given LF is highlighted. The project provides further insight on how the LPIS-LFs dataset could be used to monitor the implementation of CAP-strategic plans.

The report consists of the following main parts: (2) Data collection and sematic analysis, (3) Spatial analysis and photointerpretation, (4) Comparative analysis of the crated LF layer and the Copernicus SWF data (5) Conclusions.

1. Introduction and background

The concept of landscape features (LF) comprises the *fragments of permanent non-productive natural and semi-natural areas embedded in agricultural landscapes*. These small fragments have a key role in maintaining biodiversity and ecosystem services in the European agriculture, so they have become a priority focus of several EU policies. For example, the Biodiversity Strategy formulates goals and targets in terms of landscape features. The environmental and climate objectives of the new Common Agricultural Policy (CAP) that shall be reflected in the strategic plans of the member states, require specific indicators for monitoring and evaluating the implementation of the policy, of which one indicator is for LF. As the concept of LF is relatively abstract, which has only been recently endorsed by mainstream policy and science, there are no established quantification methods and indicators available yet. Such well-defined indicators, however, are critical for context (C.21) and impact (I.20) indicators laid out in the Performance and Monitoring Evaluation Framework (PMEF) of the new CAP.

Landscape features of agri-ecosystems providing ecosystem services and supporting biodiversity, are small fragments of directly non-productive natural or semi-natural vegetation in agricultural landscape, as well as anthropogenic structures that might also have a historical background. In most policy documents landscape features are defined as a group/list of subtypes ("features"), such as hedges, ponds, ditches, trees in line, in group or isolated, field margins, terraces, dry-stone or earth walls, vegetated areas, individual monumental trees, water streams, springs or historic canal networks. Nevertheless, there is no standard definition and typology of landscape features, and there are different interpretations in the various sectors and disciplines.

In order to facilitate interoperability of the Land parcel identification system (LPIS) and the Geospatial Application (GSA), Romania as a pilot member state was chosen. In Romania, no potential LF layer exists in IACS-GIS, only very limited number of declared GAEC/EFA LFs has been introduced, while the agricultural landscape is traditionally complex and rich of natural and semi natural features. The 4 pilot sites where LFs have been digitized and analyzed are representing complex and large variety of landscapes, while in some regions a weak ecological network is facing the challenges of large arable parcel dominance.

Based on the experiences gathered during implementing the GAEC and greening requirements of the direct payments, methodological developments of the features delineation had been elaborated, while also a testing a new simplified categorization proposed by the EC-JRC was implemented.

The challenge now is to define a LF mapping and monitoring methodology for PMEF in a context of various bottom-up GAEC/EFA CAP implementations (mainly based on VHR images by visual photo interpretation), and the new possibilities of top-down EO datasets (derived by machine learning models). The synergy of the two mapping methods seems to be the most advanced new direction to reach a complete LPIS-LF layer.

2. Data collection and sematic analysis

2.1. Review of available dataset

Romania is an example of EU countries, where majority of the agricultural area is an organically developed cultural landscape. That is why the ecological network and LF the protection is a primary goal, while the creation and restoration are not in real focus. With this background the following data governance situation has been discovered:

- Several LFs have been nominated to be preserved as part of GAEC, using detailed definition and size limits, but they are not captured into reference layer as part of LPIS.
- For greening, the list of GAEC LFs has been extended, and only the declared features have been integrated into an LPIS EFA layer using point, line and polygon geometries.
- PA keeps the full responsibility on LPIS database, but despite of the willingness of data integration among public bodies, up to now the team of the PA was not able to integrate third party data sources into the IACS-GIS system. The reason is the lack of the availability of high-resolution databases, while co-operation to reach semantic harmonization with currently developed datasets is ongoing (for example with the forestry administration).

2.2. Type and classification of Romanian landscape features

Large varieties of GAEC landscape features and buffer strips are defined in the Romanian IACS, while only minor share of the features are digitized as part of LPIS. The reason is that until year 2022 restricting the delineation to stable and declared LFs in IACS completely fulfilled the requirements. That is why several paying agencies (PA) in the EU decided to validate them against the latest ortho images based on the declarations of the farmers in GSAA. This methodology fulfils all criteria defined for LPIS layer of stable EFA and LF elements in the period of "CAP-greening" from 2015 to 2022. This approach supports the long-term protection of features declared by the farmers themselves. On the other hand, this data is not suitable to assess the status of the ecosystem services and contribution to biodiversity of a given region. Figure 2-1. shows how spatially unbalanced is the spatial distribution of the 1908 LFs, as declared by the farmers in the GSAA. 39% are field margins, 34% are ditches, 6% are AL buffer zones, 5% are wooded strips, 4% are solitary trees, while the remining 12% contains a few examples of the various other categories.

Figure 2-1 County level distribution of LFs declared in year 2022, with the example of wooded strips declared in GSAA - Romania

To implement an exact maximum of 10 or 20 meters width as a delineation rule along the entire extent of the LF is rather difficult and meaningless in case of a semi-natural arboreal features, where the width is not a constant value. As table 2-1 shows, Romania is among the member states, who had chosen to delineate the features as lines, as this seemed the least risky solution to map landscape features which are natural elements of the land cover with various width and share of wooden land covers.

A possibility of assigning a "weighting factor" and a "conversion factor" to each LF type was implemented to calculate the area of ecological interest (EFA) in relation to the greening payment. "Conversion factor" is used to estimate the area covered by the linear and point features, which should be remained as part of the system until the line and point representation remains.

Ecological Focus Area (EFA) type in IACS/LPIS (2022) [1]	Item identi-fication	Delina-tion: topology in GIS	Conve r-sion factor (C)	Weigh-ting factor W)	Area multiplicati on factor (M=CxW) [m2]	Min. width (m) or size (m2)	Max. width (m) or size (m2)
Terraces - minimum 1 meter high on slopes of over 15-16%	**TS**	LINE	2	1	2	1.5 m	3 m
Landscape elements - Hedges / wooded strips / trees in alignment	**GV**	LINE	5	2	10	5 m	10 m
Landscape elements - Isolated trees (per tree)	**AI**	point	20	1.5	30	4 m crown diameter	-
Landscape elements - group of trees / shrubs in the plain area	**GA**	polygon	-	1.5	1.5	-	3000 m2

Landscape elements - Ponds - permanent natural accumulations of standing water	**IZ**	polygon	-	1.5	1.5	100 m2	3000 m2
Landscape Elements – Ditches, including open watercourses for irrigation or drainage, no sealed sides	**RG**	LINE	5	2	10	3 m	10 m
Buffer Zones / Field Edges - Slope Protection Strips <= 12% + strips along water courses with riparian vegetation	**ZT1**	LINE	6	1.5	9	1 m	20 m
Buffer Zones and field Edges - Slope Protection Strips > 12% grassland strips along AL or cultivated with perennial grasses or perennial legumes + strips along water courses with riparian vegetation	**ZT2**	LINE	6	1.5	9	1 m	20 m
Field Margins	**MC**	LINE	6	1.5	9	1 m	20 m
Areas with short-lived forest species - Willow (Salix L) / White poplar (Populus alba) / Black poplar (Populus black) (0,1-0,5 ha)	**SF**	polygon	-	0.5	0.5	0	0
Wooded areas	**ZP**	polygon	-	1	1	0	3000 m2

Table 2-1 Rules applied in Romania to define the areas of ecological interest in GAEC and EFA, used from year 2015-2022

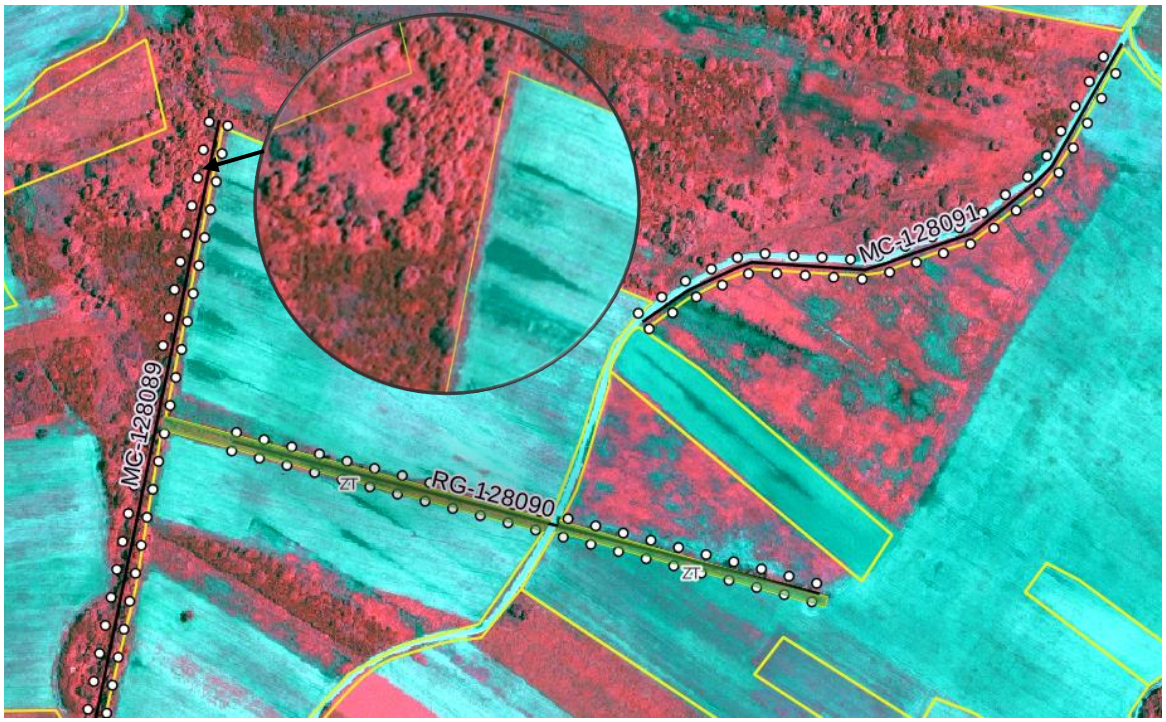


Figure 2-2 Examples presenting the advantages and disadvantages of the EFA-LF layer development as implemented till year 2022.

Examples on Figure 2-2 shows why the proof of the declared features by the administration based on a VHR image is important. The black line with white spots along represents LFs declared as line features in GSAA-2021, where MC = filed margins, and RG = ditches. Line MC-128089 do not represent a distinct feature, there is no grass margin maintained at the side of the AL, the neighbor of the AL is a continuous larger natural woody vegetation. It holds an ecological value, but does not fulfil the definition of linear LF. MC-128091 can be a grass margin, but on the current image it cannot be detected, as the parcel is weedy, not maintained. RG-128090 is a ditch, identified as dominantly grassy linear feature with <50% of trees, and the "wet" nature is not identified. Yellow = physical block boundaries.

According to the definition, buffer zones and field margins are part of the agricultural parcel as natural and semi-natural grassland or cultivated with perennial grasses or perennial leguminous plants. Buffer zones can be used to return agricultural equipment around the parcel and they cannot be cultivated for the purpose of agricultural production. Grazing or mowing is allowed, still ensuring that the LFs can be distinguished from the adjacent agricultural land.

Figure 2-3 Example of wooded strips where in certain locations along the feature the width exceeds the 10 meters limit what would mean an exclusion according to the EFA-LF definition used in greening from year 2015-2022

According to the rules until 2022 all LF must be adjacent to arable land. That is why the GSAA application is designed on a way that a line is representing the LF a wooded strips or filed margins is captured in the web-client application by the maximum distance along both side of the line, representing the maximum eligible width of the element. LFs must be directly adjacent to a declared AL parcel or to another area of ecological interest in case of a secondary LF. In the case of buffer zones and field edges that differ from those provided or protected through GAEC 1, SMR 1 and SMR 10, the surface to be qualified as an ecological focus area and calculated up to a maximum width of 20 meters. In IPA-Online, the weight and conversion factor will be automatically applied to the length of the element digitized.

Figure 2-4 Example of filed margins as defined by the Romanian Paying Agency (APIA)
Source: https://apia.org.ro/files/pages_files/GHID_INFORMATIV.pdf

Figure 2-5 Example of a field margin delineated on orthophoto images with a line using the line tool of the IPA - Online GSAA application.

As a summary, we can state that the declaration of LFs fully depended on the logic, how the EFA % can be fulfilled in a given member state. This explains, why in a LF-rich region (e.g. permanent grassland), such insignificant amount of features are declared by the Romanian farmers. LPIS-EFA layer only contains approximately 2000 features. Database with this volume will clearly not contribute to evaluating the relation of agriculture production and ecosystem services, neither supporting the maintenance or stop clearing the features away. Overploughing is a real risk, features are disappearing year to year, that is why there is a significant role of the CAP administration in this topic.

2.1. Semantic mapping between the proposed broad categories of LF, the existing GAEC/EFA-LF categorization and the corresponding feature types of the created LF layer

The Romanian LF types are mapped to a harmonized and simplified definition of LF types, which is based on functional typologies proposed by the recently published "*Landscape features in the EU Member States*" technical report [1]. The following four functional LF (FLF) classes has been defined focusing only on ecologically relevant distinctions between features representing different broad vegetation and ecological network types. Using the categories of the above-mentioned study and based on the current mapping experiences and of digitizing GAEC/EFA LFs in different EU countries the following definitions are proposed:

Woody features (A): Linear or island-like natural and semi-natural individual biotope features covered dominantly by arboreal (perennial woody) vegetation, such as trees and bushes, can be natural or planted, but never integrated part of a larger forest vegetation.

Grassy features (G): Individual patches or linear areas dominantly (> 50%) covered by permanent herbaceous vegetation, where permanent grassland dominates, embedded in an agricultural landscape, can be natural or planted, but never integrated part of a larger permanent grassland vegetation

Wet features (W): Natural or human made linear or island-like individual features, where the ecosystem – such the phenology of the vegetation – is dominantly determined by the permanent or regular presence of water.

Stony features (S): Natural or human-made appearance of individual stone features or the presence of rock layer or non-productive soil fragment on the land surface.

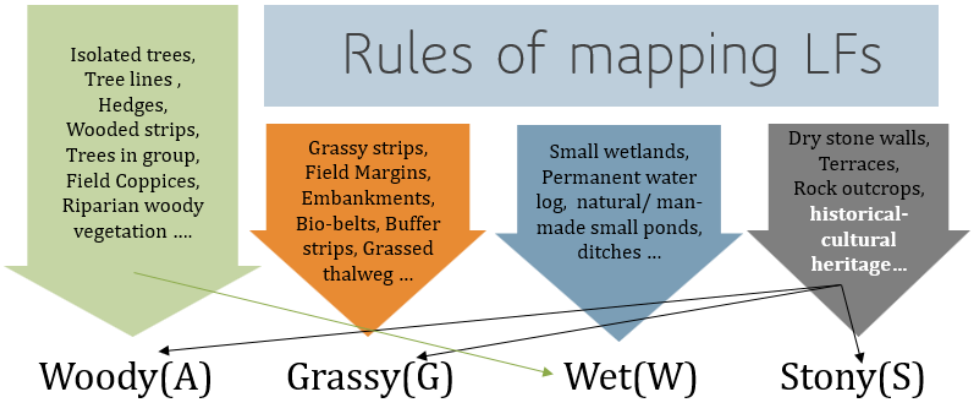


Figure 2-6: Mapping of landscape feature types to the four functional simplified LF classes

The following table presents how the LF categories defined by the Romanian PA can be mapped to the simplified functional categories as defined by the JRC:

Existing GAEC/EFA LF types in the Romanian IACS/LPIS (2022)	RO code	Simplified category	Questions identified during CAPI and based on experiences of LPIS-EFA layer generation, conditions, limits to be clarified
Terraces - minimum 1 meter high, 3 m width on slopes of over 15-16%	**TS**	**A, G, S**	Is the width an average and a maximum width along the feature?
Landscape elements - Hedges / wooded strips / trees in line	**GV**	**A**	Is an average width along the feature defined? Is any limit of canopy cover? Dominancy of arboreal phenomenon is a requirement? How linearity of the extent should bedefined to distinguish the feature from group of trees (GA)?

[1] Czúcz B, Baruth B, Angileri V, Prieto Lopez A, Terres JM,: *Landscape features in the EU Member States: A review of existing data and approaches.* EUR 31063 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-52324-6, doi:10.2760/101979, JRC128876.

			How to set the maximum width? In what circumstances does the regional defitntiton has a benefit?
Landscape elements - Isolated trees (per tree)	**AI**	**A**	How to deal with kitchen gardens and trees on AL vegetable production?
Landscape elements - group of trees / shrubs in the plain area	**GA**	**A**	Seting up rules to exclude inner non-arboreal LC. Does it have a meaning to keep the „inner-AL" and „adjancency" rules as previously detailed for GAEC/EFA?
Landscape elements - Ponds - permanent natural accumulations of standing water	**IZ**	**W**	Do wetlands belongs to here? How to define if a pound is permanent? Should a waterlog be captured as well? – Does the regularity counts? How to handle the riparian vegetation?
Landscape Elements - Ditches	**RG**	**W**	Should a minimum width be defined? Can the wet feature remain in case the presence of water is not proven? Does an automatic buffer around a linear element of water body complies with the requirements? How should the natural streams and their riperian vegetation be categorised? What should the maximum width be to keep each part of a single linear running water in the system?
Buffer Zones / Field Edges - Slope Protection Strips <= 12%	**ZT1**	**G**	Is an average width along the feature defined? Does the dominancy of the land cover phenenon (woody or grassy) matters?
Buffer Zones and field Edges - Slope Protection Strips > 12% grassland strips along AL or cultivated with perennial grasses or perennial legumes	**ZT2**	**G**	What conditions should met to consider a Buffer Zone a stable LF – as part of LPIS - and in what cases it is an agro-system services as being part of the cultivated agricultural parcel?
Field Margins	**MC**	**G**	Is an average width along the feature defined? Should dominancy of herbaceous/gramiaceae species be a requirement?
Areas with short-lived forest species - Willow (Salix L) / White poplar (Populus alba) / Black poplar (Populus black) (0,1-0,5 ha)	**SF**	**A**	Only the GSAA-declared ones, or all what can be detected?
Wooded areas	**ZP**	**A**	To map only the registered or declared ones, or all potential featured what can be detected?

Table 2-2: Mapping of Romanian LFs to the simplified functional categories

The existing LF had only been defined as located or adjacent to arable land. In this pilot we intend to test the challenges of capturing the features on the entire agricultural area, also on PG and PC.

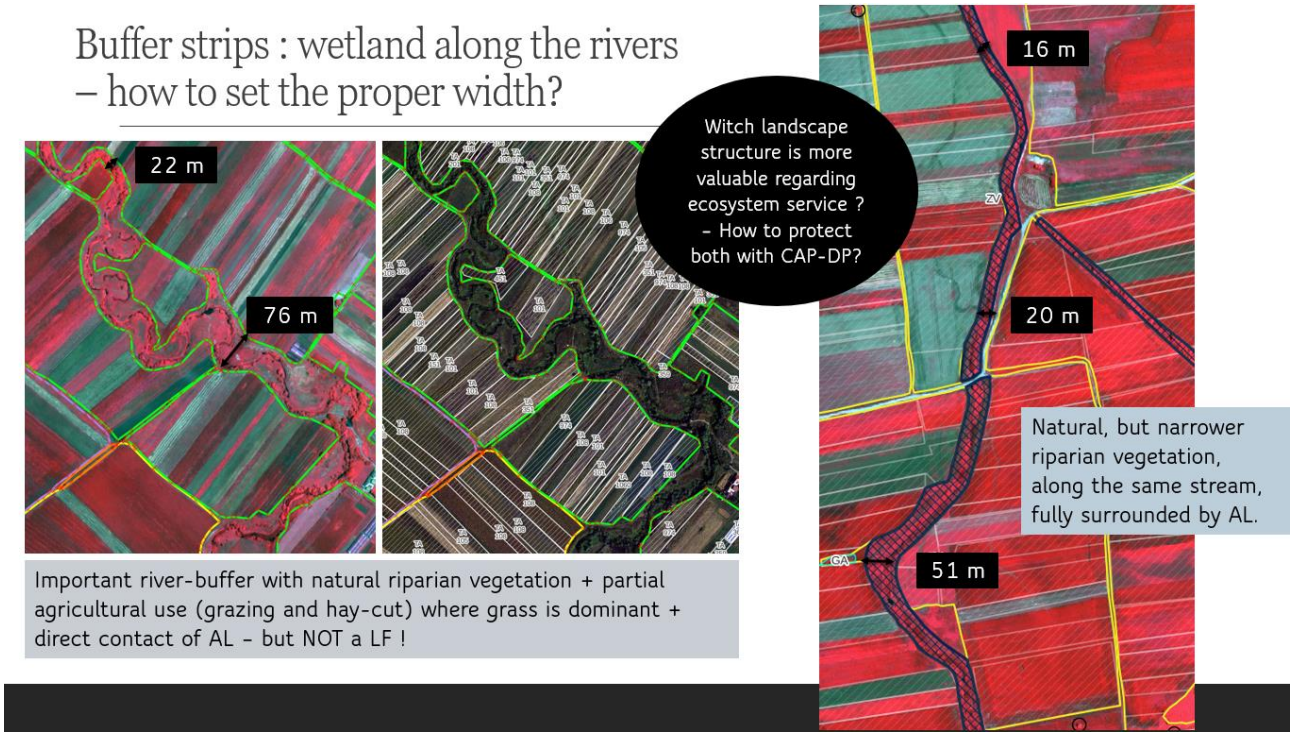


Figure 2-7: Questions analyzed regarding the categorization and delineation of natural running waters and their riparian vegetation

2.2. Definition of landscape features to digitize on the pilot sites

To be able to delineate the functional features of the agri-landscape on a logical way, resulting that the meaningful elements will be captured with their entire extent, the following main aspects had been identified, what has a direct relation to the definitions and to the mapping:

- The feature should be **independently identifiable** "by it is own", holding a distinct physical appearance in relation to the surrounding landscape during a significant period of the year. They should never bean organic part of a neighboring larger natural vegetation. Typical example of such challenge is, when a wooded strip along a ditch is bordering the parcel and a natural wetland, where similar groups of arboreal vegetation are spread. In this case the wooded strip has to be identifiable as a distinct feature from the side of the natural wetland as well, if not the trees forms part of the wetland.
- Width can only be exact if the LF boundary is captured by CAPI as the ground based **limit of the feature**, excluding the overlapping canopy extent. Image classification methods handles it differently, as the mixed pixel effect along the boundaries will always be GSD-dependent, while the size of the overlapping canopy might vary.
- An **average width** of a feature is suitable to determine the limit of maximum extent. Natural features might vary in their width, even extending the predefined maximum of 20 meters. This should be taken into consideration with appropriate flexibility, not to exclude any of the meaningful linear ecological network elements and to avoid geometrical over-separation of a functionally continuous feature. The average width should be defined according to the typical nature of a certain linear feature type, sometimes taking into consideration the local landscape structure and historical background.
- The **dominant woody or grassy land cover phenomenon (over 50% share)** in relation to the spatial extent of a given LF should be taken into account. This method itself determines the decision on the extent/length of the feature. The borderlines of the categorization of features are: (1) the dominant extent of arboreal vegetation canopy cover, and (2) the decision on the limit of an individual object. While deciding the dominant extent with CAPI the human decision is able to take into account the annual growth/invasion capacity of the local bush and tree species as well as the human contribution to the maintenance (potential cut of edges). While off-leave winter images

support the under canopy interpretation, it would lead to under estimation of leaf fall vegetation. The relative homogeneity of the woody vegetation spread will determine the extent of a feature as an independently identifiable single unit. In case of linear features ,a minimum length of a feature is recommended to be defined (30-50 meters) to avoid meaningless splitting of small units.

- **Dominancy among feature types,** in case when multiple presence of categories in a single feature is detected, the order is based on the strength of ecosystem contribution:
    1. wet (in case of continuous presence throughout the year)
    2. woody
    3. grassy
    4. stony nature.

    It means that the part followed by trees will remain a wet feature even though the water body is covered by canopy. In case of a natural habitat, where bare rocks and grass species are in mosaiced pattern, the dominancy of the grass will define the category.

The fact that the Romanian LFs has not been digitized as polygons on ortho images (only as lines) determines that their published definition do not fully follow the rules of CAPI: That is why the interpretation details and analysis of cases has been completed by the current project. A completely new LF layer was digitized on VHR images on the 3 pilot sites. A methodology of direct digitization (CAPI) of simplified LF categorization was tested integrating all the technical experiences of defining and delineating more detailed types of GAEC/EFA LFs of the previous CAP-. Following a bottom-up enlargement of LFs and from the perspective of managing the LPIS and linking the features to scheme specific requirements, a sub-division of the simplified LF categories might have been logical, while for the performance calculations both on farm and MS level the 4 simplified LF groups are appropriate.

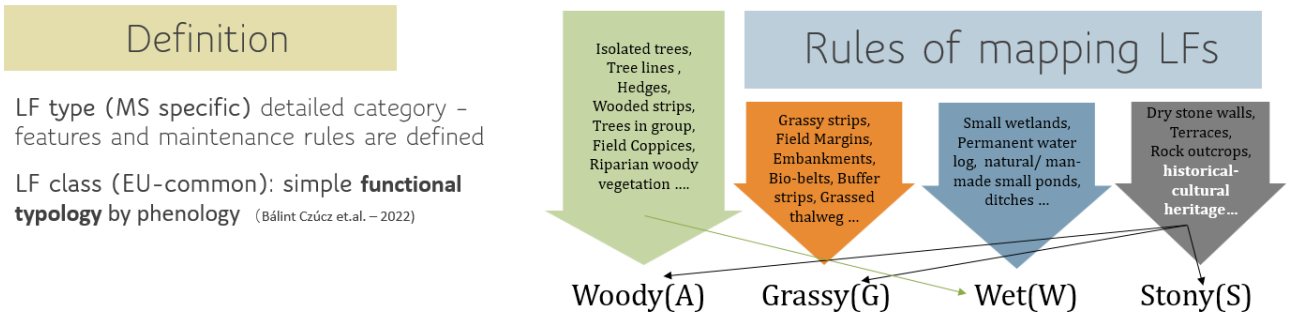
Distinction of 100% woody and 100% grassy features from those, where the dominant share of woody or grassy cover is coded will contribute to better implementation of:

- image classification/machine learning algorithms automatized validation of the features,
- prioritization of update procedures, focusing CAPI on dynamically changing elements, while the more stable ones can be validated with EO-image analyzing,
- tendency monitoring of encroachment of woody vegetation.

It hardly matters neither for biodiversity, nor for ecosystem services if there is a gap under 5 meters somewhere in the regular pattern of trees, or if a feature of irregular shape consisting of a mixture of shrubs and grass. What really matters is the total extent of the feature, and its “connection” to the surrounding semi-natural features and agricultural fields. Both can be reasonably well limited with simple LCC, width and area measures for features embedded among agricultural fields.

The categories defined according to the four functional simplified LF classes and the definitions used for digitization of features on VHR images are summarized in the following tables. Lessons learned during the multi annual development of a continuous LPIS og Hungary where all non-eligible land covers are also classified according to LF’s definitions, the knowledge organized in the EC-IMAP[2] site and the Landscape features GENERAL Fiche (Data extracted in May 2022) had all been taken into consideration.

[2] *IMAP stands for **I**ntegrated **M**odelling platform for **A**gro-economic and resource **P**olicy analysis. With IMAP the Joint Research Centre provides scientific support and tools to DG Agriculture and Rural Development for implementation, monitoring and evaluation of the CAP - https://wikis.ec.europa.eu/display/IMAP/Landscape+features_GENERAL*



Def. of agricultural landscape: LCC adjacency	AL+PG+PC: inner features + partly located within a boundary buffer of 20 m, including directly secondary LFs + **distinguishable from the neighboring fields and natural ecosystems** - Understanding the landscape structure is fundamental.
Size limits: min, max. width, area – to comply with other regulations	Max 5000 m2 and **approximate average width** of 20 m or more if suitable, focus on habitat spatial structure: skip the old rules of 2 meters gaps – only consider real artificial sealed elements Dominancy (>50%) of grassy or woody vegetation of the entire feature – not to cut a linear feature to small parts
What not: excluded as a LF:	LFs fully on yards and recreational areas, narrow PG parcels

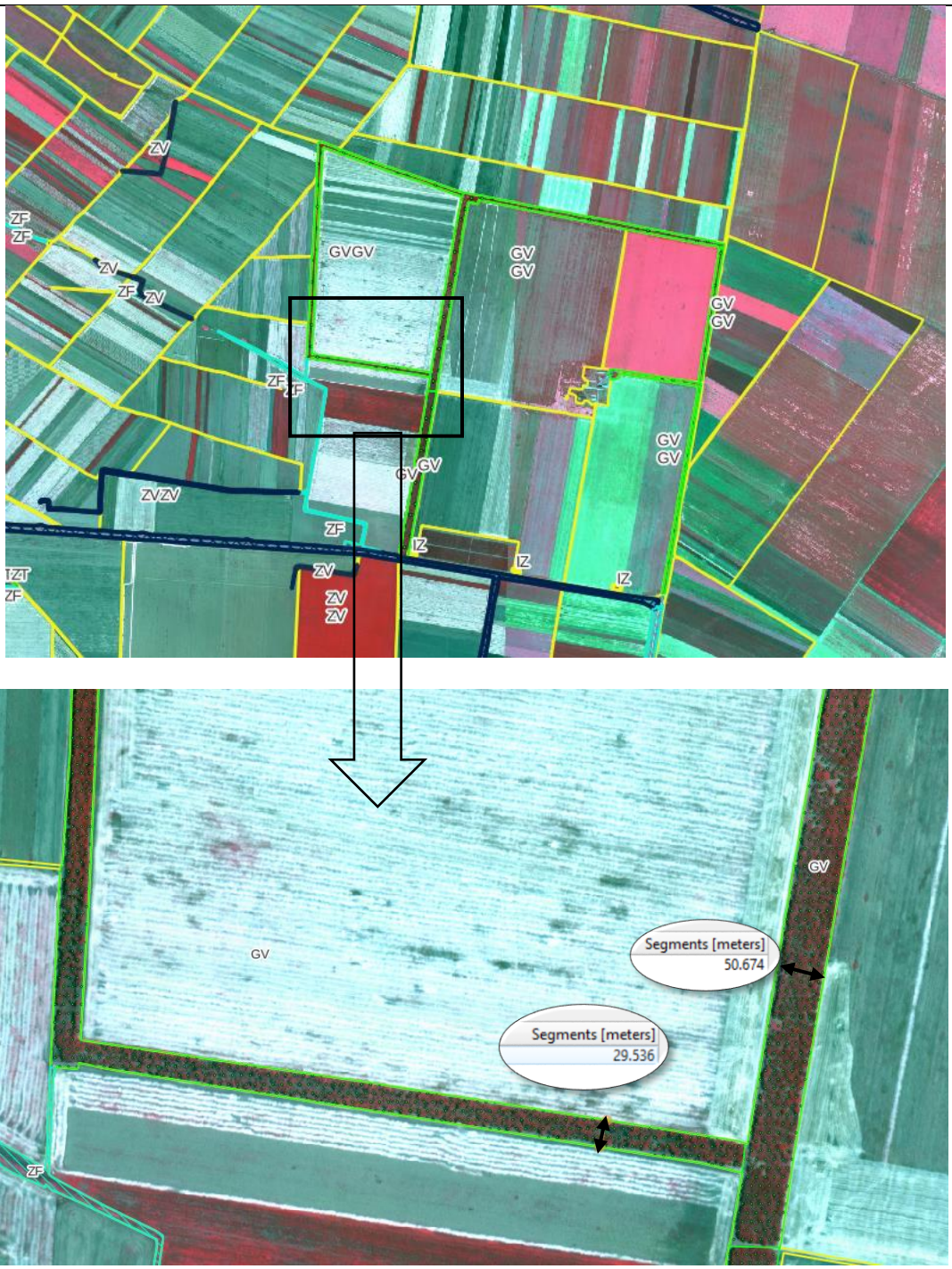
Figure 2-8 Summarizing the 4 main simplified LF categories and the main rules of the detailed definitions.

This LF categorization presents a logical subdivision of the 4 simplified categories – woody, grassy, wet, stony - according to the land cover and vegetation characteristics that can be identified and to the potential ecosystem functions. The specific size, appearance, choice of construction material (i.e., earth, stone or brick), age, land use/vegetation cover may differ across biogeographical areas.

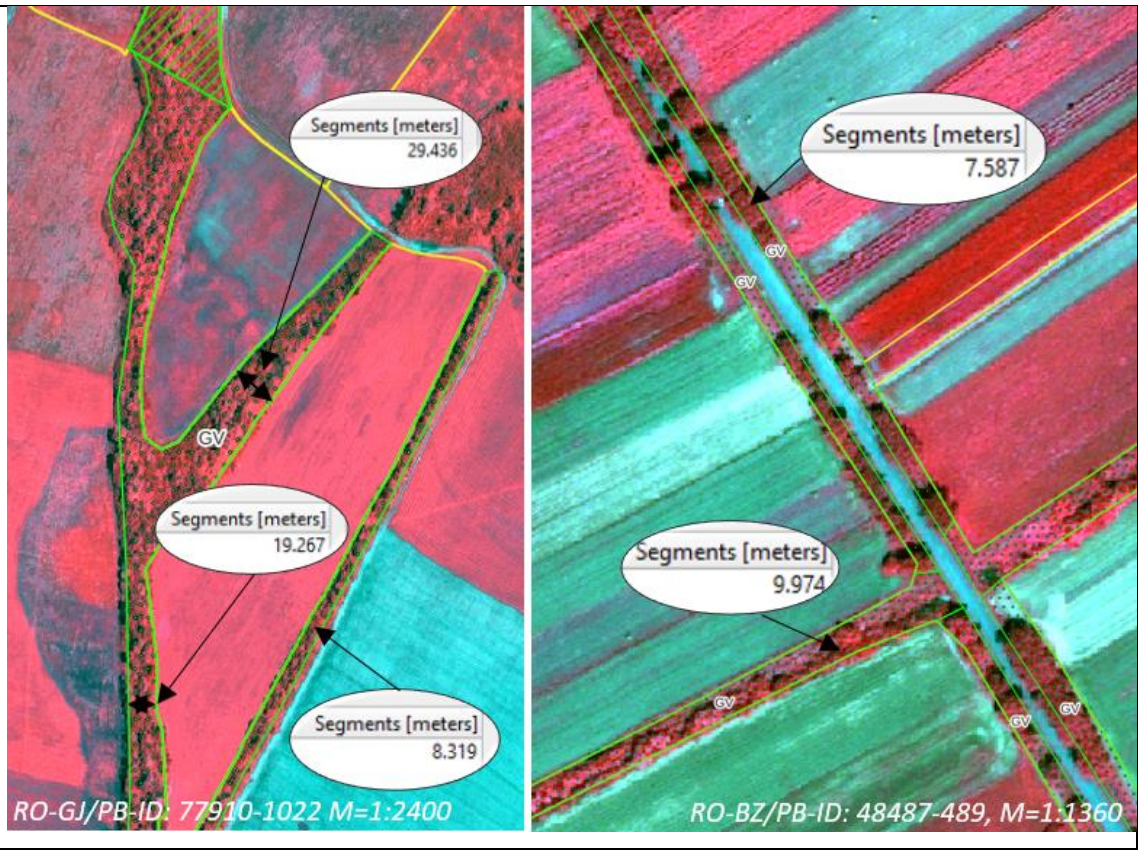
LF Type	**Woody (A)**[3]	
Definition	Linear or island-like natural and semi-natural individual biotope features covered dominantly by arboreal (perennial woody) vegetation, such as trees and bushes, can be natural or planted, but never integrated part of a larger forest vegetation.	
Sub Types	**Wooded strips, hedges/trees in line (GV)** – linear areas where woody vegetation is dominant, including trees, shrubs and herbs.	Minimum 1 meter wide and maximum 30 meters of average width in relation to the entire feature. According to local specificities, the width can be extended. Minimum twice as long as the average width. The upper limit should be harmonised with the specific rules of forestry administration, if a wooded strip is part of the forest management, prohibiting of clear cutting should be administratively coordinated (Cross-Compliance).
	Group of trees (GA) - small patchy areas embedded in an agricultural landscape where woody vegetation (including trees, shrubs and herbs) is dominant. They can range from ancient native woodland remnants to new plantations of non-native species.	Minimum size is according to the MMU of LPIS - usually 100 m2 on 1:5000 scale, maximum 3000 m2 (in case of Romania) but can be 5000 m2 - usually determined by the local rule of legal forest categorization.

[3] Codes in brackets are used during the pilot, and delivered in <RO_CODE>and <JRC_TYPE>fields of the shp files

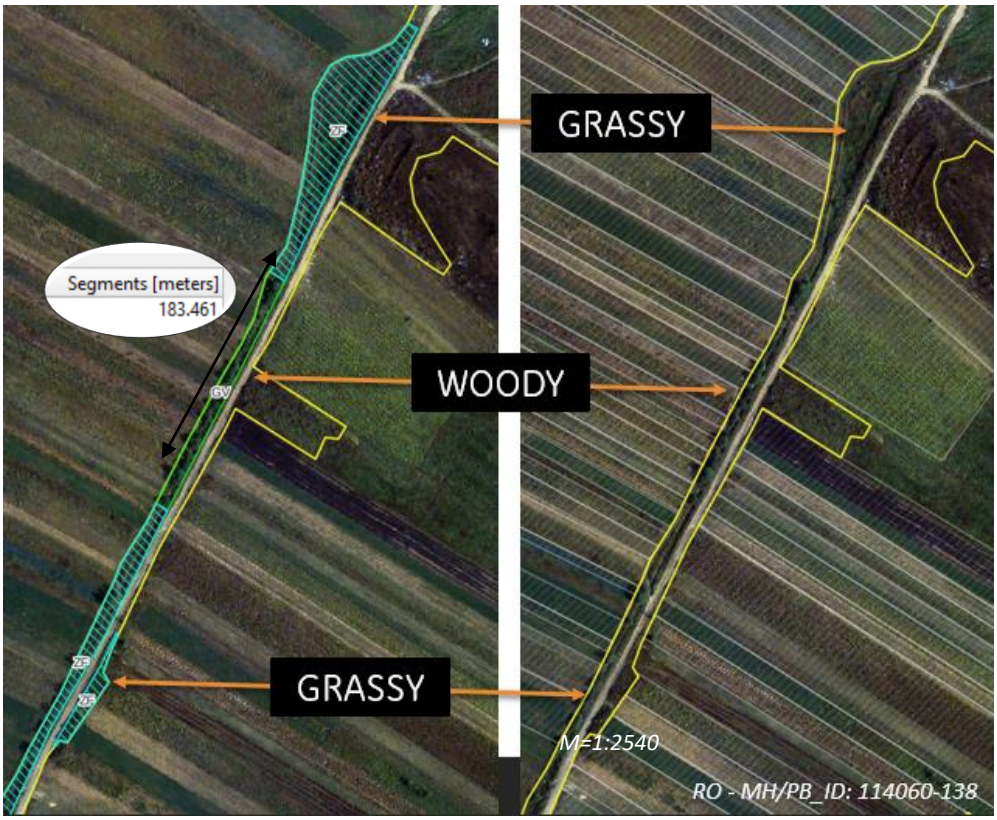
	Isolated trees (point layer) - non-productive trees occurring dispersed / scattered in croplands and/or grasslands, typically as legacies of historical vegetation and land uses	Over 4 m of diameter.
	Wooded terraces (TG) - anthropogenic structures on sloping terrains created to permit or facilitate cultivation and to reduce the risk of erosion, dominantly covered by permanent woody vegetation and flat sections that are used for agricultural production, separated by the steps.	Min 1 meter of height difference and minimum 1 meter wide, which is less or maximum the width of the flat sections used for agriculture
	Water protection buffer strip (ZV) – dominantly woody, separately delineated along water bodies	Minimum 1 meter wide and maximum 30 meters of average width in relation to the entire feature.
CAPI specialities	The actual land cover boundary is captured as the limit of the feature, excluding the canopy extent overlapping a neighboring land cover. 50% dominancy of arboreal canopy cover is considered, also in relation of bare soil, rock surfaces. Gaps of the vegetation caused by roads, ditches only count if they are artificially paved and wider than 5 meters. Any gaps, cross-roads can be taken to cut the feature (especially if they are at the edge of an LPIS reference parcel) to avoid extreme large linear elements, difficult to handle on web-map frontends (GSAA). Trees in line usually categorized as woody features based on the ratio of natural/semi natural area along the trunks and the ratio of the canopy cover. Missing trees are not taken into account, no rule of “gaps”is used, until the gras cover of the strip becomes dominant reaches the 30-50 meters length of an individual “grassy” feature. Over-split of the linear elements based on switches of dominance should be avoided, and the dominant LCC rule should be evaluated for the entire feature forming a single continuous unit. Wider grass strips with trees of smaller canopies can be categorized as grassy (G) features, based on the dominancy of the herbaceous vegetation on the entire unit. Size independent exclusion of built-in inner area and of natural herbaceous vegetation or wetland > 1000 m2. In case a running water body is not visible, neither proven by in independent database (topographical water layer or any other surway) it is taken as a wooded strip based on the tree canopy phenomenon.	
Examples	Examples of natural wooded strips with different widths, fully surrounded by intensive arable land, thus presenting an outstanding high ecological service.	



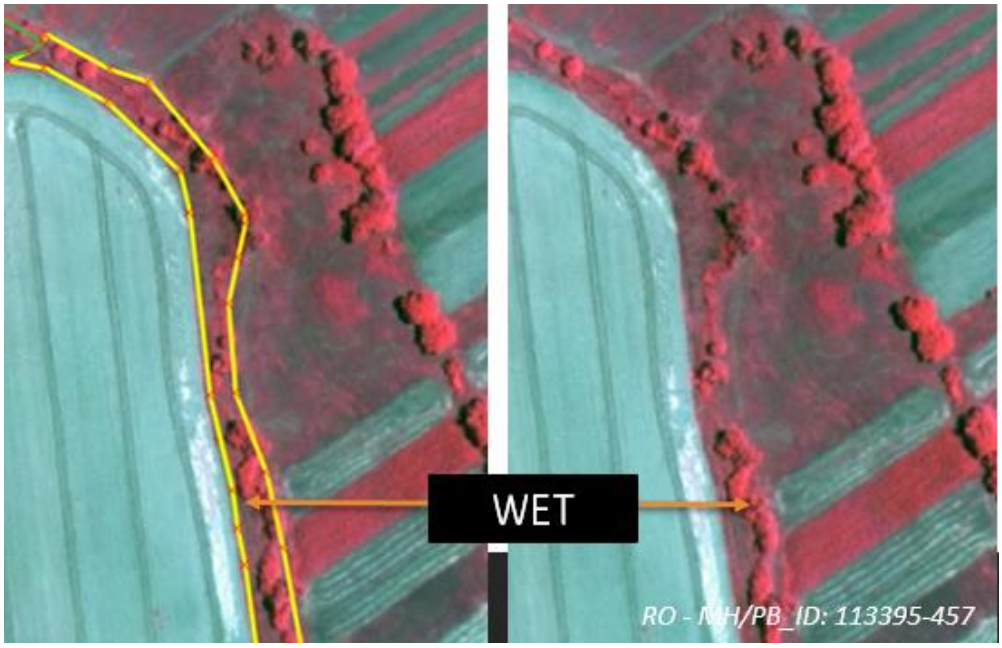
The example above shows an intensive AL region, where 30-50 m wide wooded strips form the linear biotope element in the region, but these features could not be part of any protection according to the former CAP rules applied.



LF Type	**Grassy (G)**	
Definition	Individual patches or linear areas dominantly (> 50%) covered by permanent herbaceous vegetation where permanent grassland dominates, embedded in an agricultural landscape, can be natural or planted, but never integrated part of a larger permanent grassland vegetation.	
Sub Types	**Field margins exclusively grassy (ZF) -** narrow linear areas along agricultural field borders covered dominantly by grassland	1-20/30 meters of average width in relation to the entire feature, depending on the local conditions, minimum twice as long as the average width.
	Field margins with scattered trees/bushes (ZT) – where arboreal vegetation exists, but it's density is less than 50% of the feature's area, excluding the overhanging canopy.	
	Water protection buffer strip (ZV) – dominantly grassy, separately delineated along water bodies	
	Grassy terraces (TG) - anthropogenic structures on sloping terrains created to permit or facilitate cultivation and to reduce the risk of erosion, dominantly covered by permanent grass vegetation and flat sections that are used for agricultural production, separated by the steps.	Min 1 meter of height difference and min 1 meter wide, which is less or maximum the width of the flat sections used for agriculture
CAPI specialities	If the density of arboreal vegetation exceeds 50% it is a wooded strip.	

	If the water body cannot be distinguished from the water protection buffer strip, the feature will be assigned to wet type.
Examples	 The example above shows how the meaningful split of features is possible according to the dominancy of wooden or grassy vegetation.

LF Type	**Wet (W)**	
Definition	Natural or human made linear or island-like individual features, where the ecosystem – such the phenology of the vegetation – is dominantly determined by the permanent or regular presence of water.	
Sub Types	**Ditch (RG) -** small natural streams and human-made linear surface depressions covered by water and/or riparian wetland vegetation, embedded in an agricultural landscape, including open watercourses for irrigation or drainage and/or soil erosion prevention.	1-20 / 30 m of average width, no sealed side
	Ponds (IZ) - small standing water bodies surrounded possibly by a narrow strip of wetland vegetation, embedded in an agricultural landscape	100-3000 m2 – including riparian vegetation
	Small wetlands (WL) - transiently flooded surface depressions covered by wetland vegetation and embedded in an agricultural landscape. This class includes the remnants of natural wetland or freshwater ecosystems, and human-made “constructed wetlands” created for treating wastewaters or as a refuge for species.	100-3000/5000 m2, depending on the local conditions

	Natural linear water bodies, streams together with water protection buffer strip (ZV) when the water body cannot be distinguished from the riparian vegetation, due to the lack of it's stable width, of permanent presence or of canopy cover.	1-20 / 30 m of average width, depending on the local conditions
CAPI specialities	Riparian vegetation will always be part of the feature. Ditch can be delineated together with the riparian vegetation, also with smaller and not independent units of wetlands. Dominancy of wet feature plays a role at this point, because in case a running water course or a ditch is detected, the feature is coded as a wet one, while simply based on the vegetation phenomenon it could be a wooded strip. The reason is that the presence of water directly determines the distribution of flora and fauna species. In case the ditch is > 5 meters wide and it has a clear boundary the water body can be delineated as an independent polygon, while the riparian vegetation on both sides are separate (in this case coded as ZV). If the running water body has the size and stable riverbed that it is digitized without the riparian vegetation, it is not a LF. Rules of seasonal, non-permanent water appearances should also be clarified. For example, a waterlog on AL will form a stable wetland in case it is repeated throughout several years (rules like 3 out of 5 years can be defined) and the presence of water will clearly prevent successful arable land management. On the other hand, non-permanent waterlog spots should not be delineated as wetlands they remain part of AL parcels.	
Examples	 Decision if a feature is independent or part of a larger natural ecosystem is not always evident.	

	 Example when a natural linear water body is together with water protection buffer strip (ZV), due to the invisibility of the stream in most length of the feature.

LF Type	**Stony (S)**	
Definition	Natural or human-made appearance of individual stone features or the presence of rock layer or non-productive soil fragment on the land surface.	
Sub Types	**Stone walls (SW)** rocky vertical surfaces with a variety of typologies. These long-standing anthropogenic structures are used since prehistory as retaining walls and/or as field boundaries.	0,5-3-5 m wide
	Natural elements of the ground surface (ST), can be part of the bedrock or any top-soil formations, adds ecological value with it's appearance, like salty benches on PP.	100-3000/5000 m2, depending on the local conditions
CAPI specialities	The spread appearance of stony features should not be in conflict with the PG/pro-rata methodology.	
Examples	No features had been identified on the pilot area.	

2.3. Semantic mapping between the broad categories of LF, the corresponding feature types of the created LF layer and the selected third-party dataset, the Copernicus Small Woody Features (SWF)

In Romania, currently, no external database was available to integrate into the LF layer. (There is a negotiation ongoing with the Forestry administration, but data development was at early pilot phase, the country -wide data availability cannot be planned yet.) The relation to the High Resolution Small Woody

Feature (SWF) at a cartographic scale equivalent to 1:5,000 Copernicus dataset was analyzed. The detailed description of the data used can be found here: https://land.copernicus.eu/user-corner/technical-library/hrl_lot5_d5-1_product-specification-document_i3-4_public-1.pdf

The small patchy and linear woody features (SWF +AWF) vector product of year 2015 was used, as the data of 2018 has not been published yet. The SWF product is derived by a combination of semi-automated image processing, feature extraction, classification using an Object Based Image Analysis (OBIA) approach and cloud-computing solutions, morphological analysis and manual editing and includes the following features:

Elements to be included in SWF mapping 2015	Elements to be excluded in SWF mapping 2015[1]
linear hedgerows and scrubs	stone walls,
tree rows (e.g. along field boundaries),	drainage ditches,
isolated/scattered patches of trees	grass margins,
	field boundaries without hedgerows or trees,
	any kind of 'grey' infrastructure such as roads
	artificial tree rows like olive tree plantations, vineyards and orchards[2]

Table 2-3: Thematic definition of SWF, Source: https://land.copernicus.eu/user-corner/technical-library/hrl_lot5_d5-1_product-specification-document_i3-4_public-1.pdf

All 3 available product type were used:

- **Linear SWF:** represent landscape features such as hedgerows or tree alignments that are defined by a compactness criterion less or equal to 0.75, up to 30 m width and at least 50 m length. They are only distinguished as separate attributes in the vector layer.
- **Patchy SWF:** represent areas of isolated and scattered patches of trees or scrubs defined by a compactness criterion greater than 0.75 m, at least 10 m width and with an area greater than 200 m² and less than 5,000m². They are only distinguished as separate attributes in the vector layer.
- **AWF:** Additional Woody Features that are neither linear nor patchy SWF, but which are connected to linear or patchy SWF and isolated woody features that are not linear nor patchy SWF, but which present an area above 1500m² (linear features wider than 30m, and out-of-specifications patches)

The categories of SWF data are semantically linked to the categories defined to delineate the LFs on VHR images, on a level of details suitable to manage biodiversity requirements defined in CAP implementation. Overlap with the wet feature category is possible, if only wooden (arboreal) vegetation dominates in the riparian zone of the detected water body.

Main type	Sub Type	AWF - Additional woody features	SWF - Linear structures of trees, hedges, bushes and scrub	SWF - Patchy structures of trees, hedges, bushes and scrub, if >200 m2
Woody(A)	Wooded strips, hedges/trees in line (GV)		X	
Woody(A)	Group of trees (GA)	X		X
Woody(A)	Isolated trees			X
Woody(A)	Wooded terraces (TG)		X	
Wet(W)	Ditch (RG)		X	
Wet(W)	Ponds (IZ)			X
Wet(W)	Small wetlands (WL)	X		X
Wet(W)	Natural linear water bodies, streams including water protection buffer strip (ZV)	X	X	

Table 2-4 Semantic mapping of categories defined to delineate the LFs on VHR images and of categories used in the SWF-2015 Copernicus data

The 3 sites contain the following amount of SWF-detected area. The difference in numbers of features and in size reflect very well the fundamental differences of the test sites:

Count of SWF	Sites			
Row Labels	**BZ**	**DJ**	**GJ**	**Grand Total**
Additional woody features	51	110	300	461
Linear structures of trees, hedges, bushes and scrub	309	923	1735	2967
Patchy structures of trees, hedges, bushes and scrub	44	101	230	375
Grand Total	**404**	**1134**	**2265**	**3803**

Sum of SWF_area (ha)	Sites			
Row Labels	**BZ**	**DJ**	**GJ**	**Grand Total**
Additional woody features	323,2976	40,4615	384,3308	748,0899
Linear structures of trees, hedges, bushes and scrub	40,6552	154,5082	323,8379	519,0013
Patchy structures of trees, hedges, bushes and scrub	2,5817	7,4979	19,5682	29,6478
Grand Total	**66,5344**	**202,4677**	**727,7369**	**1 296,7390**

Average size of SWFs (ha)	Sites			
Row Labels	**BZ**	**DJ**	**GJ**	**Grand Total**
Additional woody features	6,3392	0,3678	1,2811	1,6228
Linear structures of trees, hedges, bushes and scrub	0,1316	0,1674	0,1867	0,1749
Patchy structures of trees, hedges, bushes and scrub	0,0587	0,0742	0,0851	0,0791
Grand Total	**0,9073**	**0,1785**	**0,3213**	**0,3410**

Table 2-5 Area statistics of SWFs identified in the year 2015 HRL Copernicus data

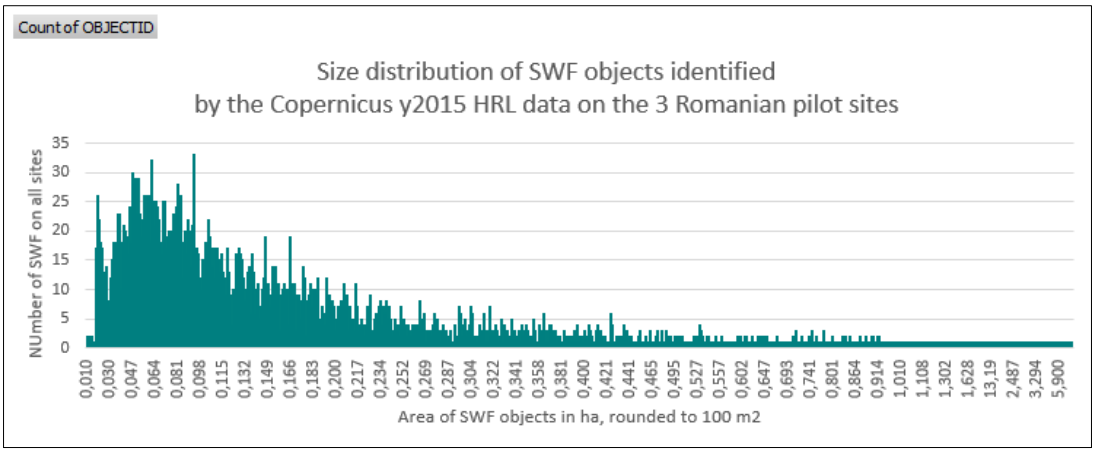


Figure 2-9 Size distribution of SWF objects identified by the Copernicus year 2015 HRL data on the 3 Romanian pilot sites

3. Spatial analysis and photointerpretation

3.1. Selection of pilot sites

Photointerpretation was done on VHR ortho images of year 2021 on 3 sites selected along the following principles:

- availability of VHR-2021 images, selection was possible among CwRS-2021 or LPISQC-2021 sites,
- finding geographical regions, representing different landscape categories, agricultural land use types and structure:

- site Buzau (BZ) is AL-dominant bringing intensive crop production both on large AL-parcels and on long narrow small parcels managed by part time farmers on very good quality soils, while appr. 25% of the physical blocks are in PG land use category,
- site Gorj (GJ) in the Carpathians in the valley of river Raul Jiu, where high complexity of traditional orchards and small PP occurs mixed with small AL parcels, this site will clearly show the weakness of the Copernicus-HRL data against the VHR,
- site Mehedinți is in the Danube valley, with several villages surrounded by traditional farming activities, including grazing on PP and maintaining traditional orchards, while AL is dominant with various parcel size and structure.

Sum of net eligible area (ha) per sites:				
LCC categories	**BZ**	**GJ**	**MH**	**Grand Total**
CP	817,14	87,24	44,97	949,35
PP	2 159,83	483,48	105,86	2 749,17
TA	7 974,16	3 505,38	9 817,24	21 296,79
Sum:	**10 951**	**4 076**	**9 968**	**24 995**

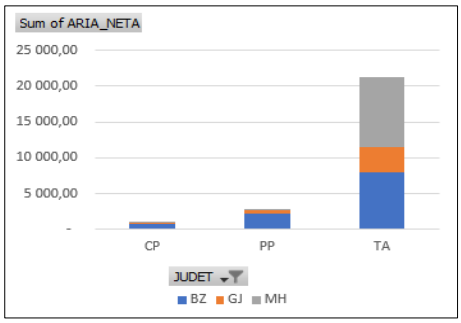


Table 3-1 Distribution of main LCC categories in LPIS on the area of the 3 Romanian pilot sites

Site	*Sensor type*	*Acq_date*
BZ_21	Worldview-2	2021.07.12
MH (DJ_21-west)	Worldview-2	2021.07.06
GJ_21	Worldview-2	2021.07.12

Figure 3-2 VHR image sources per sites

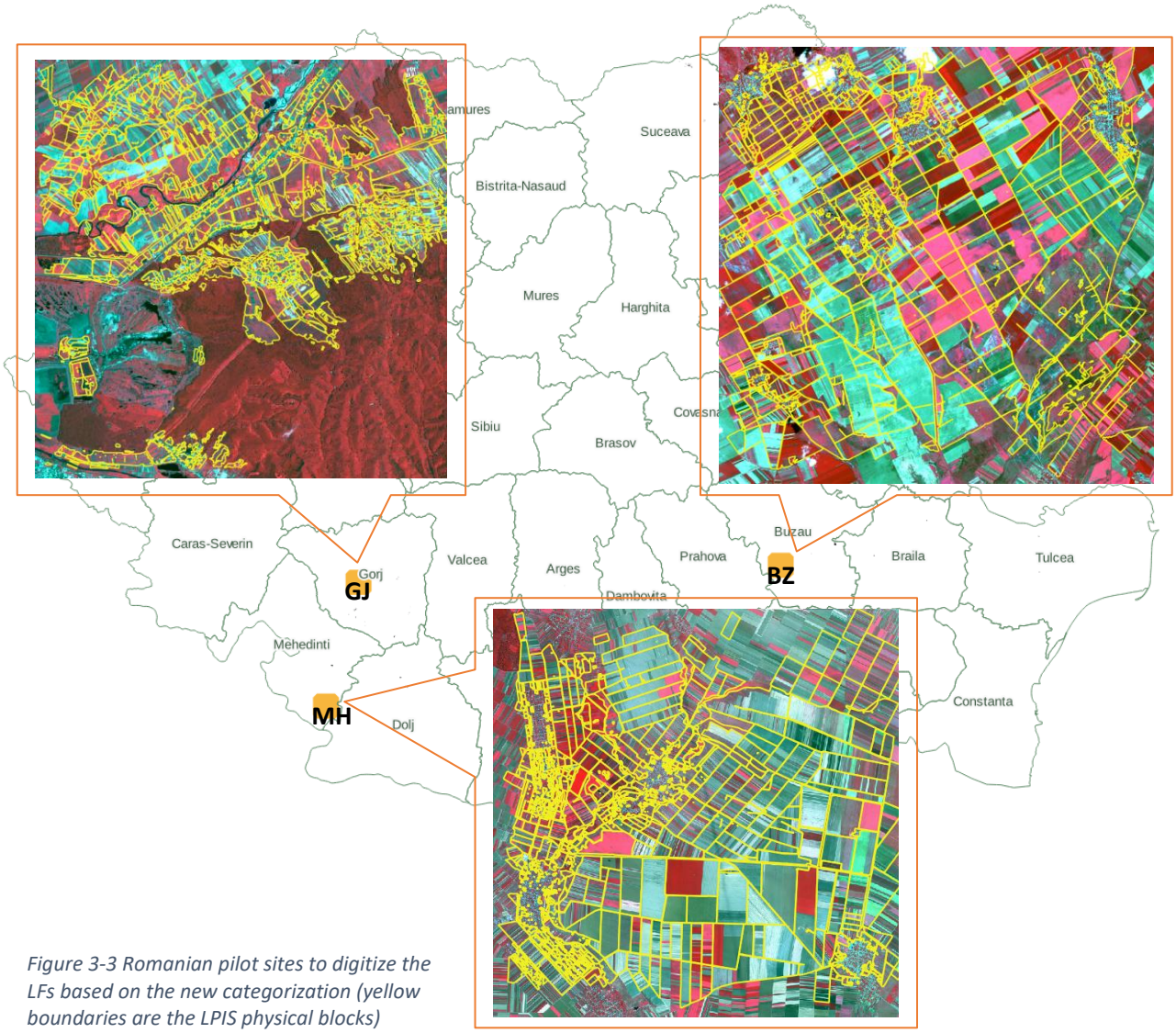


Figure 3-3 Romanian pilot sites to digitize the LFs based on the new categorization (yellow boundaries are the LPIS physical blocks)

3.2. Methodology of feature delineation

Categories and methods described in section 2.2. are applied on the Romanian pilot sites. All features have been digitized as polygons as part of the current project, except the solitary tree has a point geometry. All LFs spatially connected to AL and PG have been captured, while woody features of PCs have only been inserted into the layer if the shape and distribution of canopy clearly proved that the element is not part of the PC.

Applied rules and methods implemented to delineate the LFs on the pilot sites:

- All features adjacent to any type of agricultural parcel (AL, PG, PC and NAEA) have been captured with its entire extent (not only the adjacent part) via direct visual photo interpretation on-screen (CAPI), using QGIS SW application.
- Digitization is implemented on VHR images (GSD = 50 cm) on 1:1000 (+-500) scale, while available SPOT6-7 and Sentinel-2 images have been used to validate LCC or the stability of a feature.
- Rules and definitions set in the previous years in relation to digitizing the LPIS - GAEC/EFA layers are not entirely followed. Generalized LF types and shapes have been defined to store the main habitat types. This determines the function of a given feature, including the ability to monitor a delineated feature with automatized image classification methods. This follows the logic of AMS and prioritizes the ecological function of the entire feature using the net area, rather than applying abstract counting and weights.

- Distinct, independent features have been digitized that are not part of a larger natural/semi-natural vegetation. To be able to delineate a feature in a consistent way, the entire shape of a linear element (river and riparian vegetation) had to be investigated. Based on the landscape structure and the various appearance of the feature and its surrounding land cover elements, splitting the feature to meaningful and visible structural units was managed.
- Grassy/flowers strips as part of an AL parcel, thus part of the crop rotation and completely created as human-managed agrosystem services are not taken as LFs.
- To avoid excluding meaningful elements of the ecological network, instead of a predefined maximum width an average width is applied. Using 20 meters as a default width showed that additional analysis regarding the suitability of 30 meters seemed adequate.
- To avoid unnecessary small division of linear features, a rule of dominancy (>50% of canopy cover along the entire feature) is applied regarding the ratio of arboreal/herbaceous vegetation.
- Boundary delineation follows the rules applied in IACS-GIS/LPIS: CAPI captures the real boundary of a parcel or feature, estimating the position under the canopy or shadow. Canopy cover % is calculated for the dedicated area of the feature excluding the overlaps.
- Primary and secondary features are both considered, while maximum 5 meter wide non-paved roads can separate the agricultural parcel from the LF.

Rules and definitions tested here respect the high shape and natural ecosystem variability characterizes the natural habitats, remained in between the agricultural parcels. All natural and semi-natural linear and patchy features can become LFs as these also hold evident ecosystem services. Planted woods being stable in time are suitable as woody LFs. If these features are captured in IACS-GIS, the preservation and maintenance can be managed by the CAP interventions. Exact definition of a LF is needed in cases when the land use activities creates a non-permanent feature, which cannot always be distinguished from the productive parcel management. These are agro-system eco services, part of crop rotation and the existence and functionality is fully dependent on human land management activities, such as grassy/flowers strips on AL. If exists on the same location longer than 5 years, it becomes a PG strip, and can be a stable LF. An other example is, when a parcel with paved bed, where a stable built in wall supports the soil, but it is not forming a traditional feature, like a stone wall.

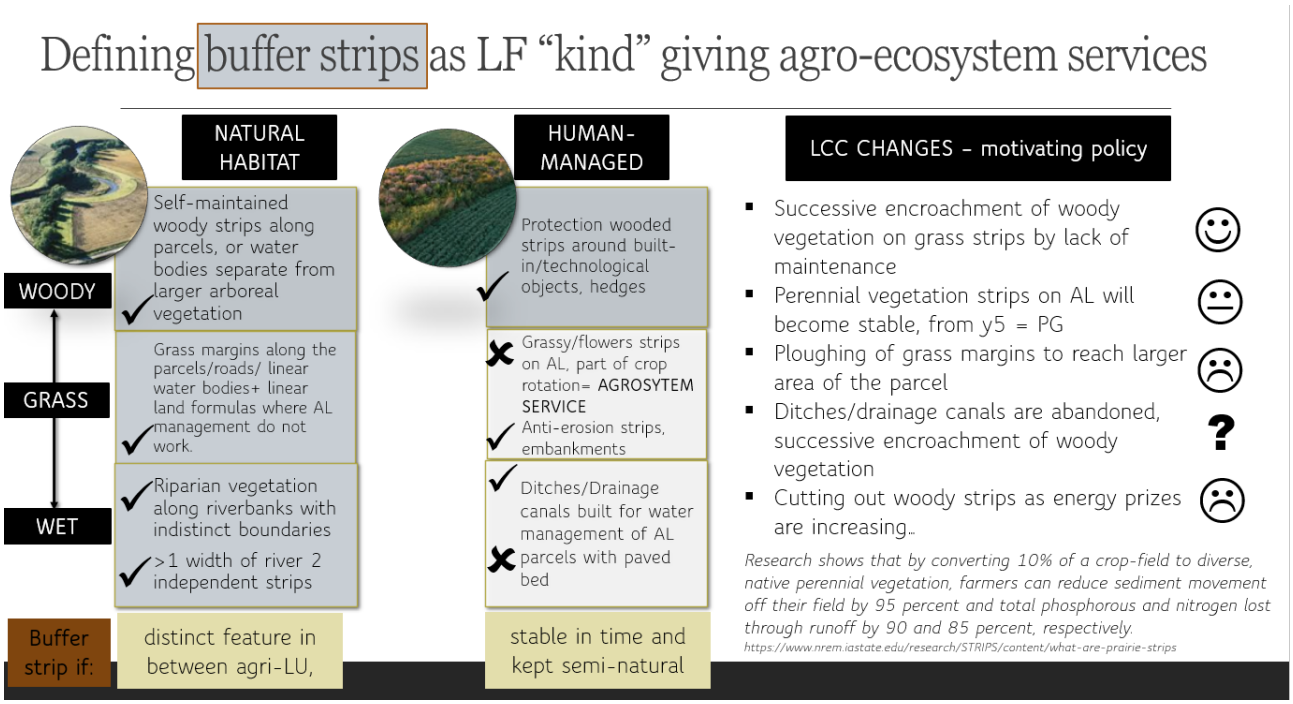


Figure 3-1 Categorization of buffer strips

Separating woody buffer strips along water bodies seems the riskiest case based on a single ortho image. Whenever the running water is not detectable under the canopies, the categorization is only possible if a topographic map is available. It was not the case in the Romanian pilot, and this would lead to different

feature classification. In case the running water body or bank was visible, the feature got the "wet" category all along it's continuity. On the other hand, woody, reed, and grassy patches are mixed in a riparian vegetation and forms a complex habitat where the arboreal/herbaceous separation is less meaningful, than the continuity and the width of the linear buffer zone. The complexity of the landscape and the dominant extent of the homogenic AL parcels will determine what maximum average width is suitable in that particular landscape. In the pilot sites the average 30 meters was applied as maximum width.

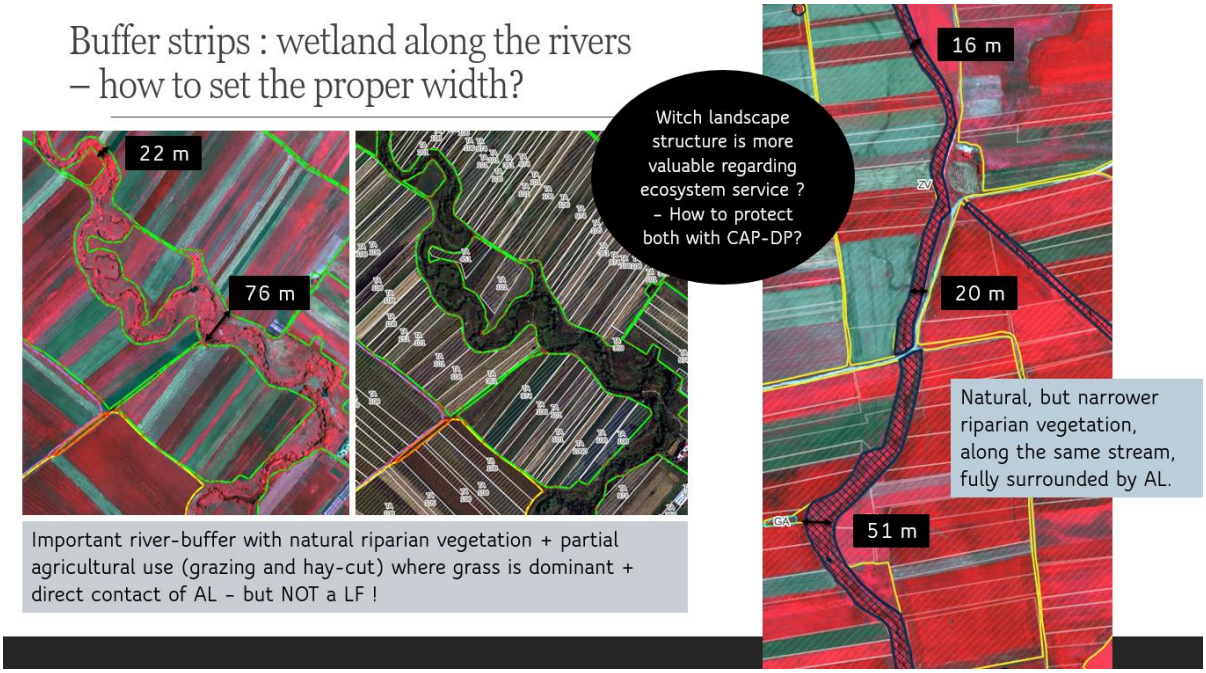


Figure 3-2 Examples of buffer strips categorized as LFs or not, according to the average width, along running water bodies on AL landscape

3.3. LF dataset produced by direct mapping based on VHR orthoimage

There were 1668 isolated trees as point and 951 polygons of landscape features digitized on the 3 Romanian pilot sites using 2 level coding. The scattered spatial distribution presents well how dominantly the type of features are related to the landscape structure, which itself is an argument for regional LF definitions and maintenance rules:

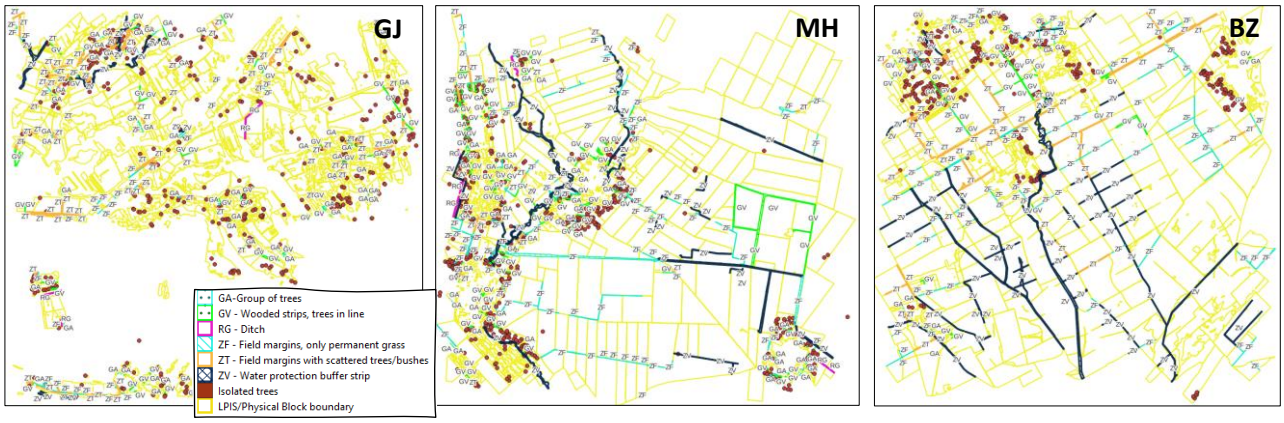


Figure 3-3 Spatial distribution of LF captured on VHR images on the 3 Romanian pilot sites

Sum of LF_area (ha)								
Detailed/Broad categories	GA – group of trees	GV – wooded strips	IZ - ponds	RG – ditch	ZF – field margin grassy	ZT - field margin with trees, bushes < 50%	ZV – buffer strips	Grand Total
Grassy					105,0573	70,1591	144,2769	319,4934
Woody	10,3581	64,0625					30,6333	105,0539
Wet			0,4031	4,6665			112,7862	117,8558
Grand Total	**10,3581**	**64,0625**	**0,4031**	**4,6665**	**105,0573**	**70,1591**	**287,6964**	**542,4031**

Count of LF_area_ha								
Detailed/Broad categories	GA – group of trees	GV – wooded strips	IZ - ponds	RG – ditch	ZF – field margin grassy	ZT - field margin with trees, bushes < 50%	ZV – buffer strips	Grand Total
Grassy					230	176	65	471
Woody	185	237					17	439
Wet			4	10			27	41
Grand Total	**185**	**237**	**4**	**10**	**230**	**176**	**109**	**951**

Area Average of LF_area_m2								
Detailed/Broad categories	GA – group of trees	GV – wooded strips	IZ - ponds	RG – ditch	ZF – field margin grassy	ZT - field margin with trees, bushes < 50%	ZV – buffer strips	Grand Total
Grassy					4568	3 986	22 196	6 783
Woody	560	2 703					18 020	2 393
Wet			1 008	4 666			41 773	28 745
Grand Total	**560**	**2 703**	**1 008**	**4 666**	**4 568**	**3 986**	**26 394**	**5 704**

Table 3-2 Area, number and average size of LFs captured as polygon on the 3 Romanian pilot sites presented in relation to detailed LF categories and the broad LF categories.

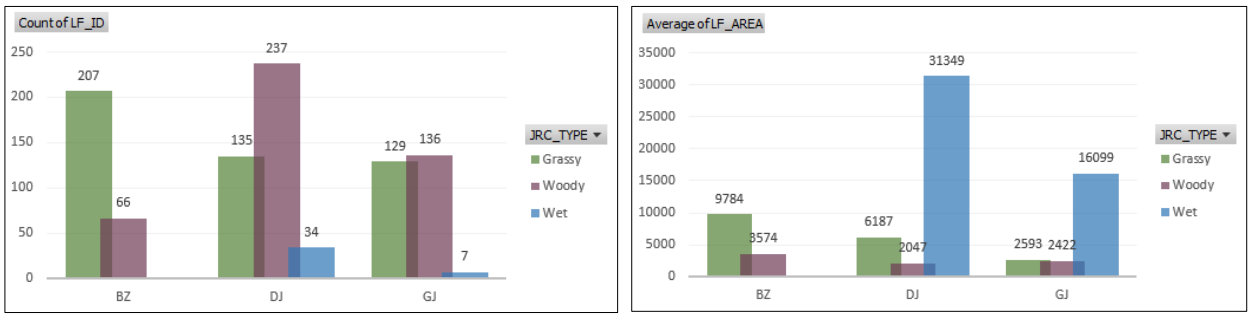


Figure 3-4 Differences of LFs types and the average size of LF types on the 3 Romanian pilot sites

The LF digitization exercise proved the increased efficiency of full extent CAPI (not just a middle-line) both in case the broad LF categories, and features delineated as compact habitats, accepting generalized size and land cover phenomenon rules. Digitization was made in autocomplete polygon mode, allowing to use the existing boundaries of LPIS, following the eligible areas of which the LF is directly adjacent to. Use of already delineated internal non-eligible areas of LPIS physical blocks was also made.

Possible feature categorization methods for developing further the LF mapping methodology, are below:

Step-1: Extracting inner exclusion geometries of the LPIS reference parcels layer under 5000 m2. If LPIS is good, as a basic rule is that all natural-semi-natural vegetation over 100 m2 or wider than 2 meters are excluded from the agri-eligible area. If PG-pro-rata is applied the limit of mapped non-eligible natural vegetation is 300 m^2.

Step-2: Overlapping the output of Step-1 with topographic and EO-derived water maps to sort the wet features. A feature is wet in case when more than 25 % of its area overlaps with the maximum extent of summarizing the following datasets (whichever exists):

- linear water bodies of topographical maps or any available water body measurements + 1,5 m buffer
- polygons of sweet water bodies of topographical maps or any available water body measurements (no buffer)
- Copernicus Land Monitoring Service – High Resolution Layer Water and Wetness (1) permanent water and (3) permanently high soil moisture data, showing the occurrence of water and wet surfaces over the period from 2009 to 2015, based on multi-temporal and multi-seasonal optical high-resolution Landsat data. Permanent water means the occurrence in at least 80% of all observations, while permanently wet areas means, reeds, peatland and wetlands being wet in at least ~60% of all observations. MMU: widths over than approx. 30 to 40m (mixed pixel phenomenon) and element over 20x20m = 400 m2 size.
- GLAD global surface water dynamics 1999-2021 data derived from all Landsat 5, 7, and 8 scenes 1999-2021 highlight the changes in surface water extent during this period. Maps include the interannual dynamics 1999-2021, discrete dynamics classes, annual water percent, mean monthly water percent, and the water percent for individual months. Water percent is calculated from only the land and water observations. All maps are publicly available within Earth Engine and for download: https://glad.umd.edu/dataset/global-surface-water-dynamics

Step-3: Define areas where AL/PG/PC eligible parcels are not adjacent, but closer to each other than 30 m, using a buffer GIS operation. These will be fully agri-dominant areas, where the linear gaps among the parcels holds fundamental ecological network elements, like wooded strips, ditches, grass margins. Delineate the LF types visually on VHR images, exclude the roads and other artificial elements from the area of the features.

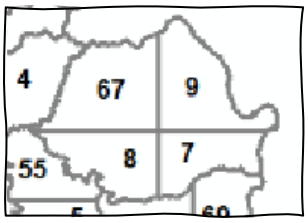
Step 4: Select LPIS RPs adjacent to SWF dataset with a single over 10 m of common boundary length, and delineate the LF visually on VHR images.

Step 5: Evaluate visually the remining boundaries of LIPS reference parcels if any LF can be categorized and delineated.

This method is faster than just digitizing the LFs adjacent to parcels one by one, and prioritizes the most vulnerable features.

4. Comparative analysis of the created LF layer and the Copernicus SWF data

4.1. Visual investigation of SWF data



We are aware of the fact that the quality of the SWF data is not equal over all EU countries. Unfortunately, the Romanian data of year 2015 do not present a very good performance. Based on a visual inspection the use of data could have been rejected. As for Romania the SWF is delivered in quadrats, 2 of the parts are represented by our pilot sites. Despite the MMU of SWF data (200 m2 and 0,75 meter width) seems to be appropriate for the LF's definition, a simple visual analysis results that the smaller features are not equally identified on the entire site. There is no clear evidence in the documentation published on the Copernicus web page, but it seems that the use of an AL-related mask or buffer hampers to identify woody features over the 200 m2 and 0,75 meter width.

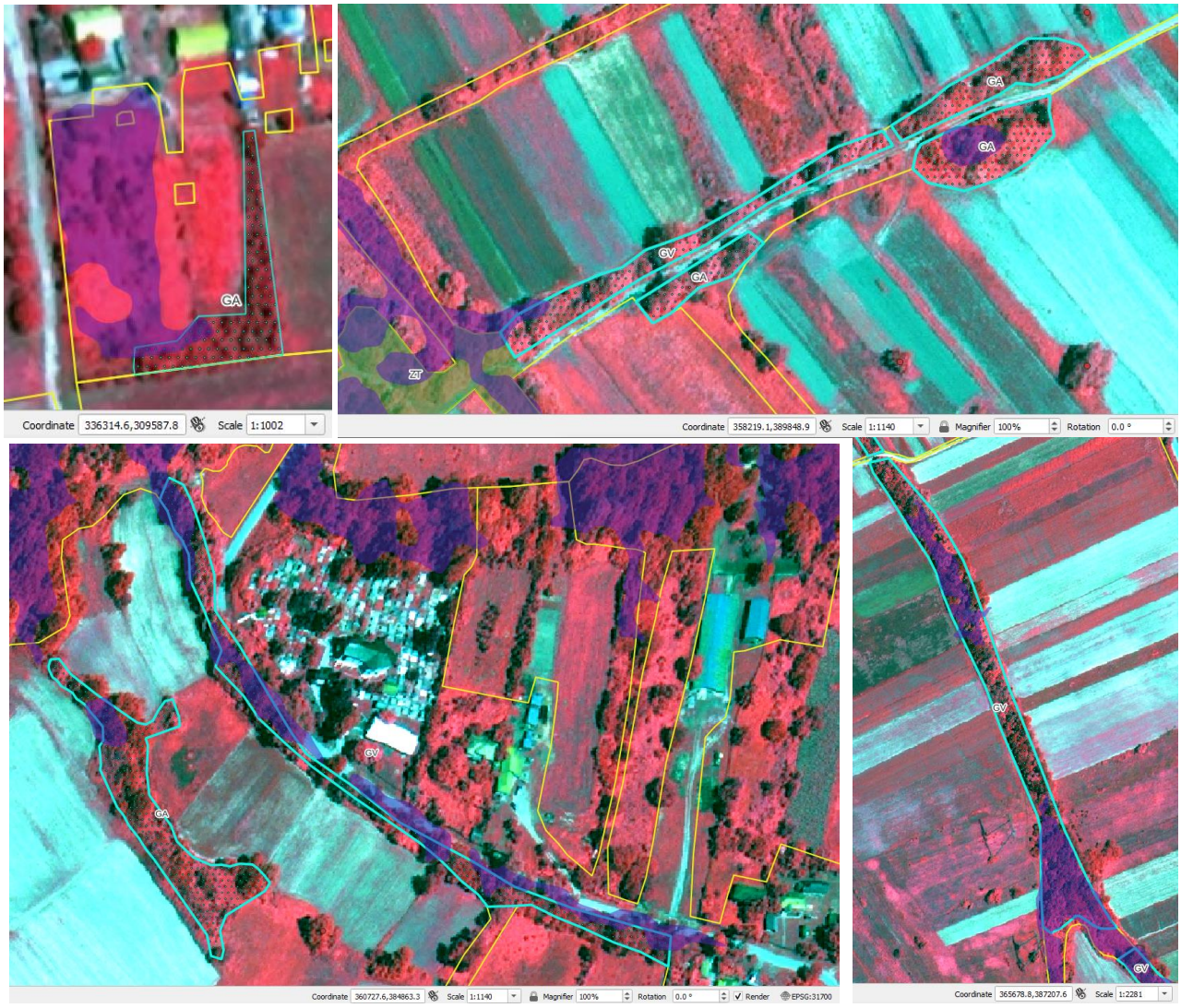
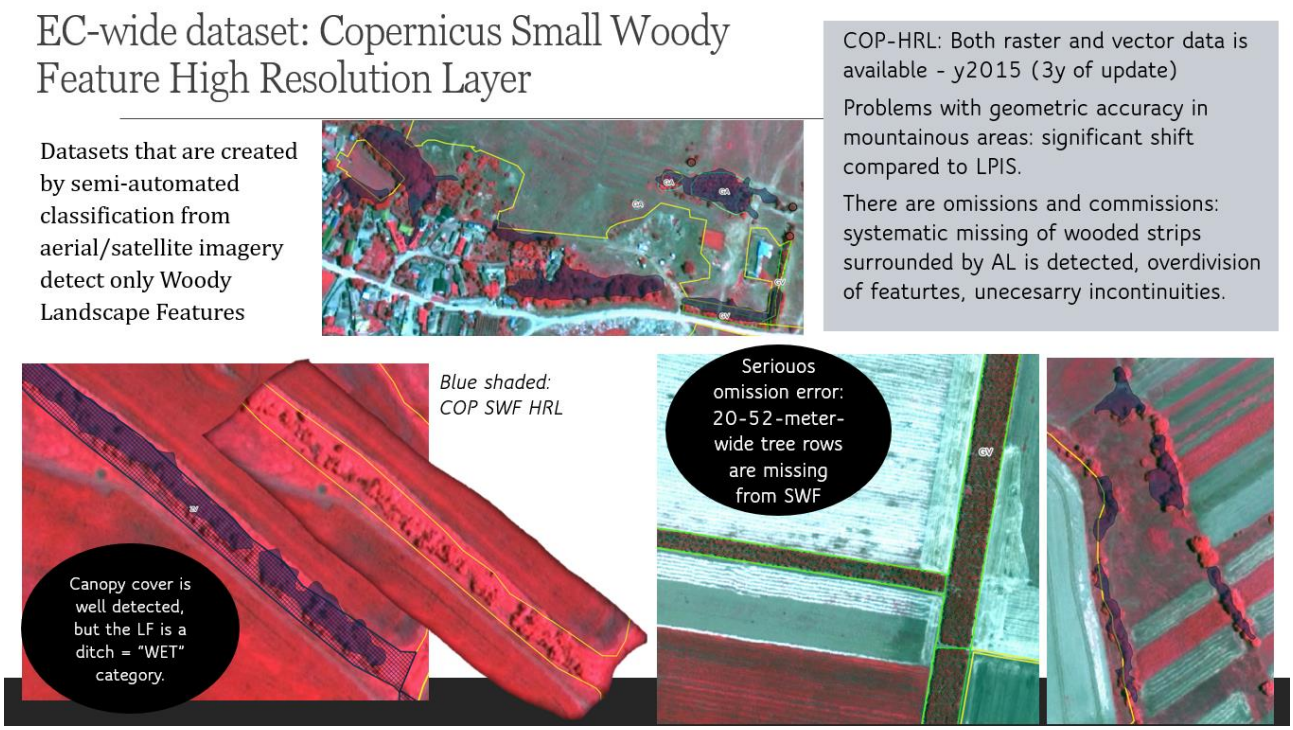


Figure 4-1 The following examples presents how scattered the SWF elements (purple shaded polygons) are, while continuous woody LFs (cyan outline with dots inside) digititsed via CAPI are not covered by the SWF features.



Another problem is that the SWF is made on image sources from year 2015, which is compared to LFs digitized on images of year 2018, in leafy period (April-June). On some areas three years could mean significant encroachment of shrubs, but the differences of the 2 datasets are mainly not depending on this factor.

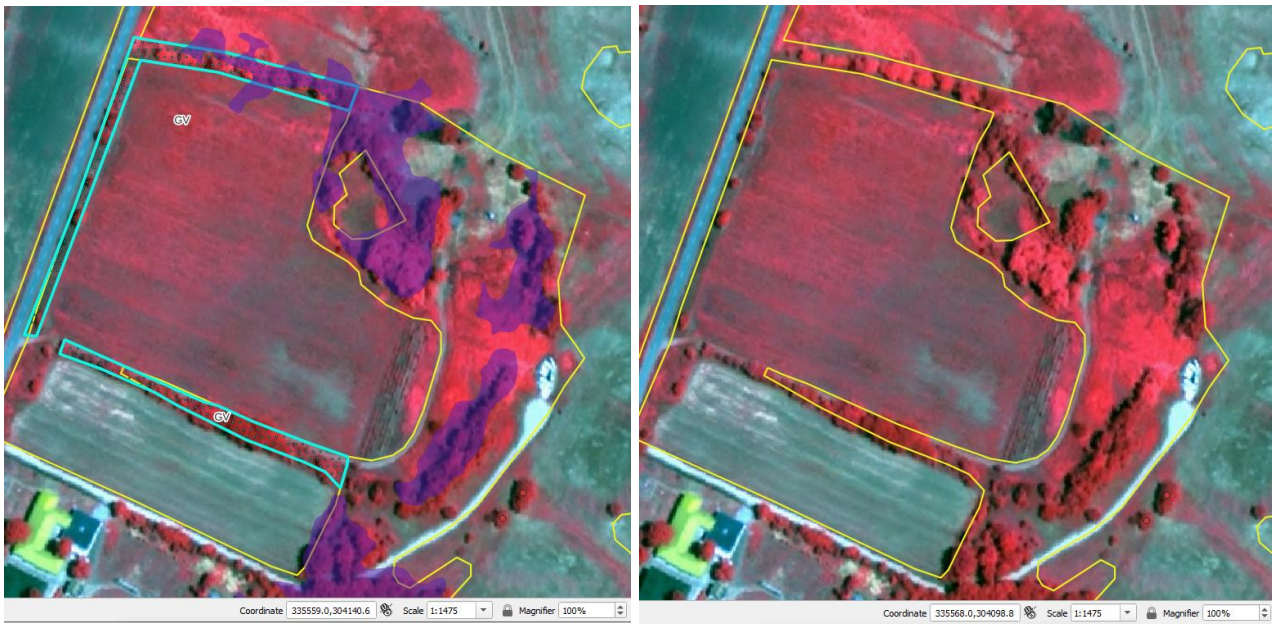


Figure 4-2 Differences of capturing wooden LF by Copernicus SWF data and by visual LF-type driven interpretation

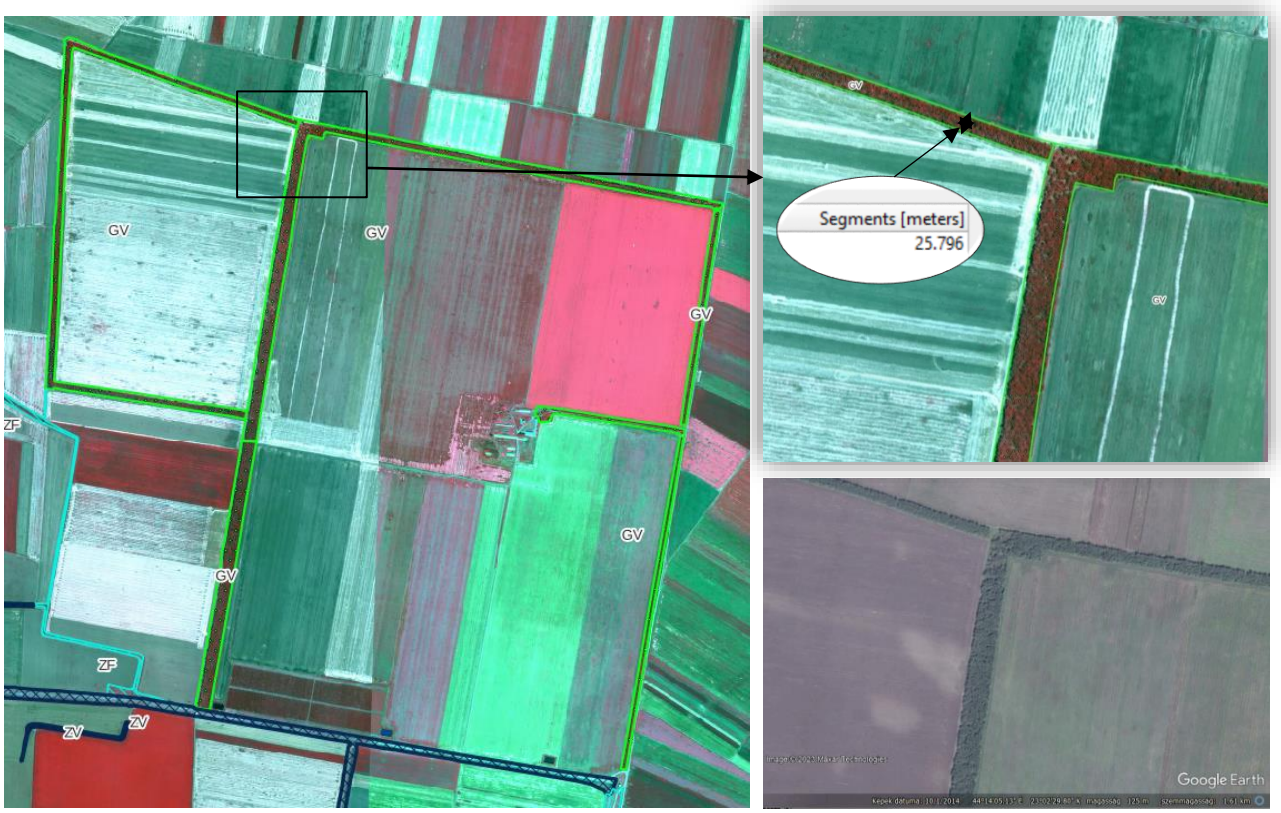


Figure 4-3 Serious omission error: in this region the dominant wooded strips are not presented in the SWF dataset at all. A Google Earth image clearly proves, that all the wooded strip had been presented in October/2014.

4.1. Statistics on matching of LF

Matching of the 2 datasets was made with simple geometrical intersect without any buffer, summing up the intersected parts of the features. SWF overlaps both with features classified as woody, and with features classified as wet or grassy. Following the definitions, the SWF overlap can be adequate, as arboreal vegetation cover under 50% of a grassy feature are identified by SWF, and riparian vegetation along water bodies were categorized as "wet" features (instead of a woody one).

LF digitization made by CAPI on an entire VHR site (GSD=50 cm) was used as reference data, to which the SWF is compared to. In all the 3 sites **23% of the woody features' area is identified by both datasets**. **15% of the area of wet features are detected by the SWF data**.

Code	*Sub types used in the RO pilot*	*Type=broad LF categories*	*Sum area of LFs delineated by CAPI (m2)*	*Area geometrically overlapping (intersect/m2)*	*Area % of LFs identified by SWF data*
GA	**group of trees**	Woody	103 581	35 141	33,93
GV	**wooded strips**	Woody	640 625	104 895	16,37
IZ	**ponds**	Wet	4 031	261	6,47
RG	**ditch**	Wet	46 665	5 089	10,91
ZF	**field margin only grassy**	Grassy	1 050 573	6 928	0,66
ZT	**field margin with trees, bushes < 50%**	Grassy	701 591	64 770	9,23
ZV	**buffer strips**	Wet	1 127 862	176 596	15,66
ZV	**buffer strips**	Woody	306 333	102 633	33,50
ZV	**buffer strips**	Grassy	1 442 769	41 193	2,86
Sum:			**5 424 031**	**537 507**	**9,91**

Table 4-1 Summarized area of geometrically overlap between the LF layer and the Copernicus SWF dataset divided by the LF sub types used in the RO pilot and by the recommended broad LF categories

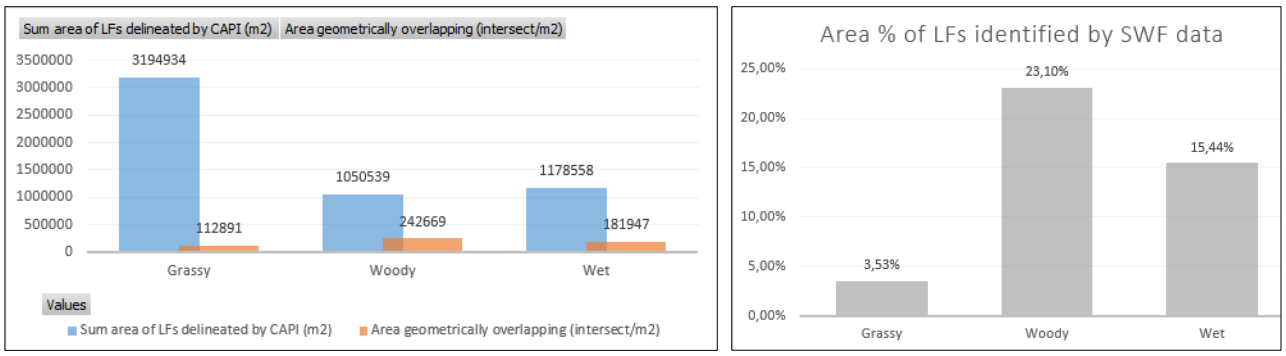


Figure 4-4 Sum area and area ratio of features where both datasets are matching, ie. LFs are identified by the Copernicus SWF data

4.2. Accuracy measures to compare the two datasets

A simple confusion matrix presents the findings whether the two woody LF delineation methods are commonly mislabeling each other. In our case, the diagonal elements represent the woody LF elements for which the SWF data is equal to the LFs digitized via CAPI on VHR images, while off-diagonal elements are those that are mislabeled by the SWF dataset.

Summarized results for the 3 sites is the following:

PREDICTED: Copernicus SWF dataset (2015)	ACTUAL LFs captured via CAPI on 3 RO pilot sites: Positive	ACTUAL LFs captured via CAPI on 3 RO pilot sites: Negative
Positive	**TP** Woody LF is identified in both dataset = visually interpret LFs and the SWF overlaps (intersect). 53,7507 ha Nr of distinct LFs: 268	**FP - Type I error** Woody element exist in SWF, but not in the woody categories of LF data. 626,9107 ha Nr. of SWF: 1774
Negative	**FN – Type II error** Digitized as woody LF, but not presented in the SWF data. 57,0742 ha Nr. of features: 249	**TN** Not presented in the SWF data and not a woody feature in the LF dataset, but grassy (410) or wet (24) features. 316,7077 ha Nr. of features: 434

True Positive (TP) is the number of correct predictions that an example is positive which means positive class correctly identified as positive.

False Negative (FN) is the number of incorrect predictions that an example is negative which means positive class incorrectly identified as negative.

False positive (FP) is the number of incorrect predictions that an example is positive which means negative class incorrectly identified as positive.

True Negative (TN) is the number of correct predictions that an example is negative which means negative class correctly identified as negative.

Figure 4-5 Confusion matrix to derive accuracy measures comparing the 2 data collection methods

Precision will present what proportion of positive identifications of the SWF dataset was actually correct:

$$\text{Precisison} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{53{,}8}{53{,}8 + 626{,}9} = 7{,}9\ \%$$

Recall or sensitivity is a measure of how woody features were labeled as woody by the SWF dataset. Also called as “true positive rate”, meaning what proportion of actual positives was identified correctly:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{53{,}8}{53{,}8 + 57{,}1} = 48{,}51\ \%$$

The false positive rate (fall-out) is the probability of falsely rejecting the null hypothesis:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} = \frac{626{,}9}{626{,}9 + 316{,}7} = 66{,}44\ \%$$

In this respect the area overlap (intersect) of the 2 datasets shows the correct identification of the SWF existence, what determines the type of the feature as woody, but not the correct delineation of the feature's entire extent. This means, that the accuracy measures will describe the identification of the woody vegetation inside a given feature, while the extent of the feature itself cannot be derived from the SWF dataset.

In the confusion matrix only the 951 digitized LF polygons have been evaluated against all the SWF features which are adjacent to LPIS eligible area within a buffer of 2 meters. The 1668 points of single trees are not presented in the statistics.

The numeric results strengthen the visual impression that the SWF data is not suitable for direct LF delineation. The accuracy measures reflects that the SWF data is not usable to delineate functional features of the ecological network. SWF data is generally suitable to detect the presence of woody vegetation, while some narrower linear strips surrounded by AL seems missing. One of the main reasons is that for LF delineation a functional definition of a continuous habitat is implemented. During LF CAPI features surrounded by other land uses and visible boundaries are delineated and classified based on the vegetation phenomenon (woody/grassy/dominancy), ecological conditions (wet) and size limits. Different size limits will determine the potential function of the feature inside the ecological network: linear elements, patchy island elements etc. Thus, the content of the SWF is just the arboreal (woody) vegetation type, what can contribute to a LF as the main vegetation category or also as a minor type.

There seems to be 2 main logical directions to use the SWF for new CAP implementation and monitoring, which are not fully independent from each other:

1. Keeping the administration of LFs in IACS-GIS, developing the data on VHR image sources, continuing to fulfil the 1:5000 mapping scale and 100 m^2/2m linear MMU as required in LPIS. In this case, meaningful monitoring of how the agri-related ecosystem functions and biodiversity is changing will only be possible if all the LFs related to agricultural land were digitally mapped. A positive effect is, that complete LF mapping IACS-GIS itself strongly supports the conservation of existing LFs , independently from the fact wether the feature is disponible and declared by the farmer. Further advantage is that already mapped features can be offered for the farmers in GSAA avoiding LFs validation before the payments of a given year. This concept will ensure the adequate input data to derive indicators for CAP to monitor ecosystem services (SO6), preserving landscape features, habitats and species (R34, R.31). Copernicus SWF data can be integrated at the following data development and data management steps:
    a. Validating the encroachment of arboreal (woody) vegetation on the area of the already delineated features, and updating the woody feature categorization according to the dominancy rule.
    b. Automatic verification of existence and enlargements of woody features, and only assigning a woody feature to CAPI in case there is any change of the extent is foreseen.

    Integration of external data sources into LPIS-LF database is a challenge, opens serious data interoperability and data management issues. If the concept of mapping the entire set of LFs is not chosen by a MS, as a minimum requirement, the EFA-layer needs to be included.

2. Deriving indicators for the CAP monitoring system on the base of other continuous datasets than IACS and LPIS. Most likely, the requirements of the refresh time will determine remote sensing-based datasets, where the annual or 3-4 yearly update can be managed with the same image classification method, ensuring the type and spread of errors to be similar. A possible solution cloud be the Copernicus SWF data. The weak point of such external (IACS-independent) dataset is, that features on agricultural landscape < 300 m^2 are not entirely detected, while these forms the most vulnerable elements of the agri-ecosystem. Apart from the known weaknesses, analyzing the change of SWF in relation of LCC and land use combinations stored in LPIS would give clear indication of trends.

5. Conclusion

The following main conclusions are found as a result of the Romanian pilot exercise:

- The broad categories of LFs types (woody/grassy/wet/stony), as proposed by the JRC, are fully adequate, and presents enough details to detect the changes of the agroecosystem services, while reduces the burden on sophisticated feature categorization.
- The LF should be defined based on it's entire extent, with categorization of land cover dominancy. To store the share of woody/grassy information on feature level is valuable for updates based on EO-image sources. Although the 1:5000 scale on-screen visual mapping of LFs cannot be substituted, the thematic categorization can be based on ML-derived EO products.
- The best approach is to create a geometrically fully suitable LF layer as part of the LPIS data, extending the LFs adjacent to AL/PG/PC parcels, as this supports the management of elements in IACS and PMEF monitoring solutions at the same time.
- In case the full coverage of potential LFs are not chosen to be maintained, technically the presence of woody, wet and grassy features can be monitored based on the collection of 3rd party datasets – such as any national datasets, specific image classification results and also the Copernicus HRLs - based on the agricultural parcels adjacency and size.
- Unfortunately, the Copernicus HRL data did not comply with the expectations because of multiple reasons: (1) the data delineates the woody vegetation, but not the extent of the LF that is why semi-natural features partially covered by woody vegetation cannot be entirely detected, (2) serious omission of smaller features in agri-dominant zones and the lack of annual data updates. This latter issue could be further improved and homogenized throughout Europe.

References

[1] INSTRUCŢIUNI DE COMPLETARE A FORMULARULUI DE CERERE UNICĂ DE PLATĂ 2022 INFORMAŢII PRIVIND SCHEMELE DE PLATĂ/MĂSURILE DE SPRIJIN - INSTRUCTIONS FOR COMPLETING THE SINGLE PAYMENT REQUEST FORM 2022 INFORMATION ON PAYMENT SCHEMES/SUPPORT MEASURES
https://apia.org.ro/despre-apia/scheme-de-plata-masuri-de-sprijin-derulate-de-apia/

Technical Guidelines on IACS spatial data sharing. Part 1. Data discovery (2020).
https://marswiki.jrc.ec.europa.eu/wikicap/images/f/f8/DS-CDP-2019-04_REV2_TechnicalGuideline_IacsSpatialDataSharing_Part1_Final.pdf

Methodological report on resolving interoperability issues in reusing IACS data in LULUCF
https://marswiki.jrc.ec.europa.eu/wikicap/index.php/File:Methodological_Report_on_Resolving_Interoperability_in_Reusing_IACS_Data_in_LULUCF.pdf

Data Interoperability Pilot: Analyzing the usability of generalized GSAA data in context of crop classification
https://marswiki.jrc.ec.europa.eu/wikicap/index.php/File:Classif_pilot_Report_JRC_CsD_final.pdf

The INSPIRE data models (https://inspire.ec.europa.eu/Data-Models/Data-Specifications/2892)

The INSPIRE data specifications (https://inspire.ec.europa.eu/dataspecifications/2892

European Environment Agency Copernicus Land Monitoring Service – High Resolution Layer Small Woody Features – 2015 reference year Product Specifications & User Guidelines, Tobias LANGANKE, version 3.7 of 2020-02-26
https://land.copernicus.eu/user-corner/technical-library/hrl_lot5_d5-1_product-specification-document_i3-4_public-1.pdf

IMAP stands for ***I***ntegrated ***M***odelling platform for ***A***gro-economic and resource ***P***olicy analysis. With IMAP the Joint Research Centre provides scientific support and tools to DG Agriculture and Rural Development for implementation, monitoring and evaluation of the CAP - https://wikis.ec.europa.eu/display/IMAP/Landscape+features_GENERAL

GUIDANCE DOCUMENT ON THE LAND PARCEL IDENTIFICATION SYSTEM (LPIS) UNDER ART. 5, 9 AND 10 OF COMMISSION DELEGATED REGULATION (EU) NO 640/2014 AND ON THE ESTABLISHMENT OF THE EFA-LAYER REFERRED TO IN ART. 70(2) OF REGULATION (EU) NO 1306/2013 CLAIM YEAR 2018 ONWARDS
https://marswiki.jrc.ec.europa.eu/wikicap/images/a/ae/DS-CDP-2018-11-LPIS-guidelines2018-clean.pdf

Amy H. Pickens, Matthew C. Hansen, Matthew Hancher, Stephen V. Stehman, Alexandra Tyukavina, Peter Potapov, Byron Marroquin, Zainab Sherani, Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series, Remote Sensing of Environment, Volume 243, 2020, 111792, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2020.111792 (https://www.sciencedirect.com/science/article/pii/S0034425720301620)

List of Abbreviations and definitions

AL	arable land
APIA	Agenției de Plăți și Intervenție pentru Agricultură, Agricultural and Intervention Agency - the Romanian Paying Agency:
AWF	Additional Woody Features
CAP	Common Agricultural Policy
CAPI	Computer Aided Photo Interpretation
CwRS	Control with Remote Sensing
EEA	European Environment Agency
ESA	European Space Agency
EO	Earth Observation
GAEC	Good Agricultural and Environmental Conditions
GIS	Geographical Information System
GLAD	Global Land Analysis and Discovery
GSAA	Geospatial Aid Application
GSD	Ground Sampling Distance
HRL	High Resolution Layer
IACS	Integrated Administration and Control System of direct payments
LCC	Land Cover Class
LPIS	Land Parcel Identification System
LUCAS	Land Use/Cover Area Frame Survey
LU	land use
LF-EFA	Landscape Features categorised as Ecological Focus Area (CAP requirement from 2015-2022)
LF	Landscape Features
MMU	Minimum mapping unit
ML	
MoA	Ministry of Agriculture
MoE	Ministry of Environment
NAEA	Non agriculture eligible area
PA	Paying Agency
PB	physical block = reference parcel type of LPIS
PC	Permanent crop
PG	Permanent grassland
QC	Quality Control
RO	Romania/Romanian
SMR	Statutory Management Requirements
SWF	Small Woody Features
VHR	Very High Resolution satellite images, with sub-meter spatial resolution

List of figures

List of tables