Analytical Review for Consolidating Habitat Mapping Within Europe: Insights from Portugal, Spain, and France

Author: Christopher Jose Pintado Sacaquirin (ChristopherJose.Pintado@autonoma.cat)

Directors: Alba Brobia [a], Dr. Joan Masó [a]

[a] Grumets Research Group, CREAF, Edifici C, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain

Abstract:

Habitat mapping plays a significant role for biodiversity conservation and environmental policy implementation. However, even with progress over the years in this field, there is still the need to make classification systems comparable across countries for harmonized maps at European scale. This paper addresses the challenges of standardizing habitat classification across three conterminous European countries by presenting a comparative analysis of habitat data for Portugal, Spain and France with focus on the cross-border regions between these countries. We have developed a prototype of harmonized regional legends according to Annex 1 of the Habitats Directive codes using a three-step approach involving data compilation, adaptability with GIS software like QGIS and MiraMon, and legend unification. Our work consists of two methodologies that either incorporate a Sites of Community Importance (SCI), or habitat displayed approach based on their data's relevant value. Diverse datasets have been integrated into coherent habitat maps, despite challenges related to data accessibility and opposing standardizations formats across countries. The results demonstrate improved processes in development and adaptation for unified habitat mapping, also emphasizing the need for improved data sharing and cooperation among institutions for better results. This study contributes to the broader goal of standardized habitat mapping across Europe, providing a framework that can be expanded and refined to include broader horizons.

Key Words: Habitat mapping, Europe, Annex 1 of HD, EUNIS, GIS

1. Introduction

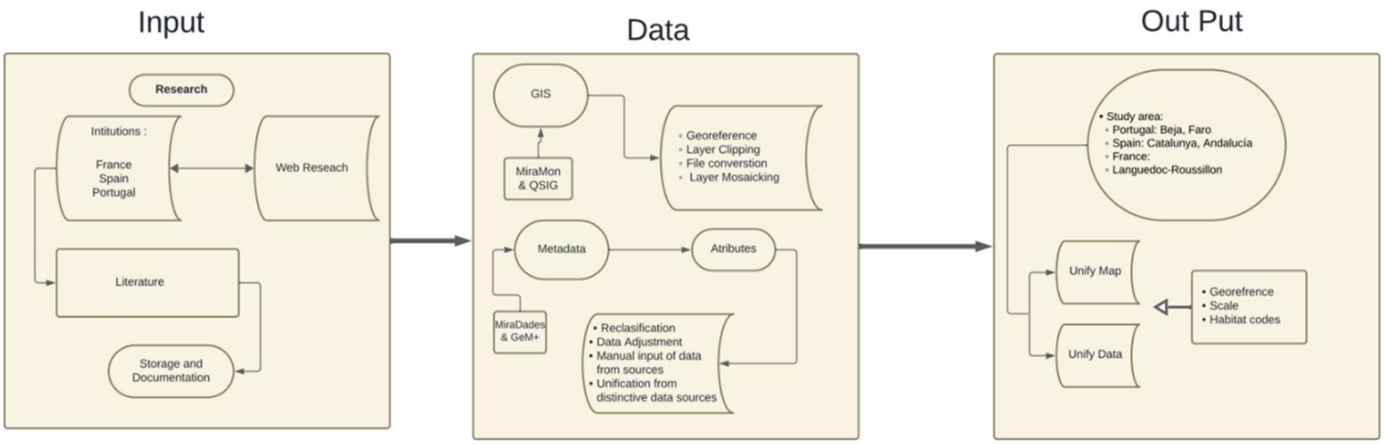
Habitats are the natural environments where organisms reside, characterized by interactions between living organisms (biotic components) and their surrounding physical conditions (abiotic components). Accurate habitat mapping illustrates the spatial distribution of different habitats within a specific region, is a fundamental tool in environmental policy and biodiversity conservation efforts. These maps are pivotal for developing, implementing, and monitoring biodiversity conservation strategies, such as the European Union's Biodiversity Strategy for 2030 and the Nature Restoration Law[1].

Since the 1992 European Union Habitat Directive (92/43/EEC) [2] represents a key piece of European environmental legislation to promote the maintenance of biodiversity by requiring EU Member States to take measures to ensure the conservation of endangered species and habitat types. It established the first steps to create a network of natural species and habitats that not only encompass each individual country but Europe as a whole. This directive has resulted in significant advancements in habitat mapping incentive across Europe.

Despite growing development, challenges remain, particularly in achieving compatibility among different classifications standards used across Europe. Some notable mention are the Annex 1 of the Habitats Directive (HD), which list natural habitat types of community interest, and the European Nature Information System (EUNIS) [3], which is a comprehensive pan-European system that provides a unified approach to habitats mapping developed by the European Environment Agency (EEA). These two classification systems adhere to the principles and methods of habitat conservation outlining a system that carters to European nations environmental needs, thus standard approaches are needed to properly generate habitat mapping at a continental scale.

This paper attempts to tackle this challenge by presenting a comparative study approach, of three selected conterminous European countries habitat maps, being that of Spain (Catalunya) with France (Languedoc-Roussillon), and Portugal (Beja & Faro) with Spain (Andalusia), with the aim to harmonize each regional legend according to Annex I of HD and/or the EUNIS habitat codes.

Figure 1. Research workflow



Thus, the experiment presents a framework for reclassifying distinctive natural habitats across four different regions in three countries. This reclassification is achieved through the homogeneous harmonization of legends and metadata. A three-step approach was developed to accomplish it (Figure 1).

Firstly, data was compiled from various sources, including governmental and regional entities, as well as universities contributing to the topic. This dataset was documented in an Excel sheet, noting each source's usability. Efforts were made to contact these institutions directly, though these attempts were not always successful. In parallel, the study was enhanced by conducting a literature review of previous research on different classification methodologies. This provided a more enriched approach and insight into the subsequent implementation of the study.

Stage two focused on adapting the dataset for processing and eventual display on a homogenized habitat map. Geographic Information Systems (GIS) software such as MiraMon[5] and QSIG[4] were used to adapt the current layers to the work objectives. An in-depth data analysis was conducted to adjust the dataset, ensuring it reflected the project's purpose of displaying either the Annex I of the Habitats Directive (HD) and/or the European Nature Information System (EUNIS) habitat codes.

Finally, the processed data was used to unify and harmonize the legend and habitat map. This final stage provided insight into a unified process by which different sets of habitat codes could be interpreted and implemented using GIS.

In retrospect there has been a movement to properly standardize the classification approaches around the entire continent of Europe. In field surveys on the ground (in-situ data) are the first source of information for habitat map creation. The literature tells us that current in-situ have been enriched and partially validated with remote sensing data though Earth Observation (EO) approaches, like using high-resolution (VHR) satellites data, and via automatic processes from a former set of data set as in the study by Nicolau et al[9]. One of the references to have taken place and resonates the most with this EU habitat mapping work is the theoretical study by Slootweg [10], which transforms habitat codes like CORINE and land cover map classes of the Stockholm Environmental Institution (SEI) and into EUNIS codes by method of translation based on previous dictionary set of equable set of information.

This study takes inspiration from this literature. It explores the initial thought that harmonizing different assets of data can be feasible, from previous metadata information that may be available. It suggests the theoretical possibility of transforming current set of data from previous identical set like using ANNEX 1 code to be transpose into EUNIS outlined by Slootweg et al methodology[10]. Even though gathering has been promising it highlights a significant gap in the existing case studies and theoretical articles do not address reclassification across international borders and their complexities, focusing instead on specific and homogeneous zones within their respective countries. To be coherent about the process in which this study influences the following Table 1 presents a research synthesis review.

Table 1: Studies on Habitat Classification.

References	Detail of study	Solution methodology	Application type
Tomaselli et al. [6]	Investigation of the effectiveness of FAO-LCCS taxonomy to fill the gap between LC and habitat domains when mapping LC classes to habitats.	A qualitative and quantitative analysis using Earth Observation (EO) data	Theroretical Study
Volpaton et al.[7]	Mixed methodological framework to assess and map biodiversity HNV agricultural landscape using habitat quality as a proxy.	Data collection using Rapid Assessment Cards and data obtained from an expert knowledge-based model.	Case study in Burren, Ireland
Adamo et al. [8]	Knowledge-based solution to automatic habitat mapping, using only expert rules and EO data.	Used of high-resolution (VHR) satellite data and LIDAR data. LC classes are labelled according to the LCCS taxonomy, which offers a framework to integrate EO data within situ and ancillary data.	Case study in the south of Italy
Nicolau et al. [9]	Harmonizing historical vector categorical maps with related modern maps.	The approach aims at the correction of geometric distortions and semantic disagreements using alignment processes and analysis of thematic coherence	Case study in Portugal
Slootweg et al.[10]	Study aims to harmonize land cover maps than can be used for all European applications under the LRTAP Convention.	A reclassification using CORINE and SEI land cover codes to be translated into EUNIS classification system.	Theroretical Study

The aim of this study is not only to unify different datasets into a coherent habitat map but also to foster an intellectual understanding of the various classification systems across national borders according the NUTS2 administrative level. Achieving this required in-depth empirical research to identify the most appropriate data, and a detailed analysis of the metadata since they vary significantly.

This project also seeks to address the challenges of developing a unified method for displaying a comprehensive map that encompasses all four regions, aiming to have an integrative map unified not only by their habitat codes but also their scale/resolution, geographic reference systems, and attributes table structure that facilitates the semantic interpretability that consist of outline area coverage and habitat descriptive labels.

2. Materials and Methods

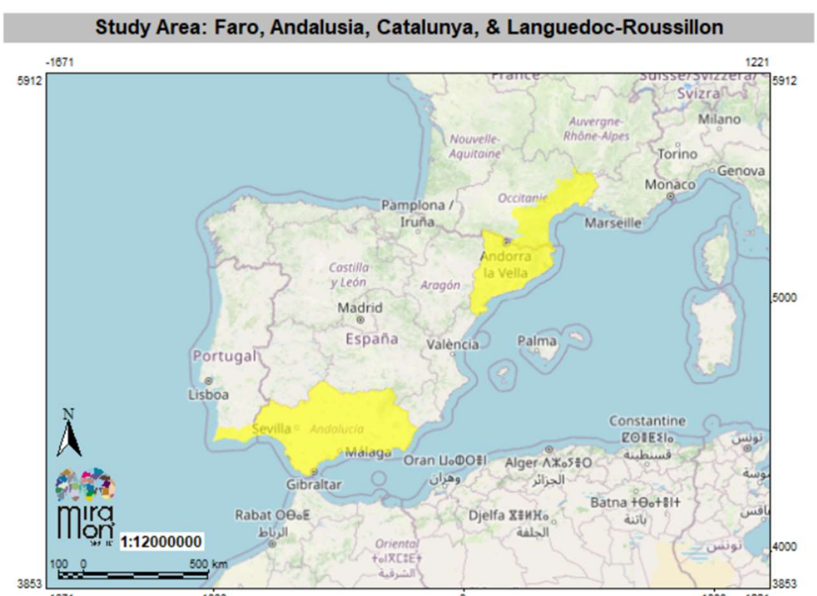


Figure 2. Study Area

i. Study Area

The study area is defined within the broader framework of the Nomenclature of Territorial Units for Statistic, Level 2 (NUTS 2) which is a classification system developed by the European Union to establish a geocode standard for referencing the subdivisions of countries for statistical purpose.

To the following European regions have been selected: Faro (Portugal), Andalusia (Spain), Catalunya (Spain), and Languedoc-Roussillon (France) (Figure 2). The selection criteria is grounded on their geographical position, as their proximity permits the analysis between conterminous nations' boundaries and related habitat maps.

3. Data Collection and Metadata Documentation

Field	Explanation
ID	
Member country	
Geographic scope	according to INSPIRE:EU, National, Regional, Local
Geographic extent	Name of region of locality of geographic scope is not EU or national
Dataset name original	
Group	Thematic groups of datasets
Short description / Resumé	Short description of of the dataset
Data provider	
Reference year(s) available	
Geometry	Raster/Vector (Point/Line/Polygon)
Format	Shape, Geopackage, TIFF
CRS/EPSG	Coordinate System
Spatial Resolution/Scale	
Thematic coverage	wall-to-wall or selected ecosystem types
Temporal extent	
Time series available	
Time series steps	
License	based on CIS2 CodeList
License	
Attribution & use restrictions	
URL	URL for data and metadata
Reference/Technical documentation	if available
Brief description by data provider	
Contributor	Organisation/Partner who created the entry
Additional Comments	
Type of Habitat	National codes, Annex I HD, CORINE, etc.…

Figure 3. Data documentation

The crucial aspect of this study was the recompilation of data. The process of obtaining this data presented several challenges, including linguistic differences, varied data formats, and issues related to institutional ownership and public accessibility versus intellectual property. To address these challenges, we have developed a standard format of documentation in which different data sets were collected and evaluated. The template used for this documentation assesses different metadata categories (Figure 3).

This categorical specification is used to narrow down the metadata attributes each available data source for easy analysis and comprehension. Its facilitated the procedure of narrowing down and selecting which data groups to use as part of the methodological test of unifying attributes to one another.

Table 2. FAIR principles

FAIR principle:		
Findability	F1	(Meta) data are assigned globally unique and persistent identifiers
	F2	Data are described with rich metadata
	F3	Metadata clearly and explicitly include the identifier of the data they describe
	F4	(Meta)data are registered or indexed in a searchable resource
Accessibility	A1	(Meta)data are retrievable by their identifier using a standardized communication protocol
	A1.1	The protocol is open, free and universally implementable
	A1.2	The protocol allows for an authentication and authorization procedure where necessary
	A2	Metadata should be accessible even when the data is no longer available
Interoperability	I1	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
	I2	(Meta)data use vocabularies that follow the FAIR principles
	I3	(Meta)data include qualified references to other (meta)data
Reuse	R1	(Meta)data are richly described with a plurality of accurate and relevant attributes
	R1.1	(Meta)data are released with a clear and accessible data usage license
	R1.2	(Meta)data are associated with detailed provenance

The metadata template used also assesses fairness of the habitat data using the following criteria found in Table 2.

The FAIR Guiding Principles for scientific data management and stewardship [18] are designed to improve the supporting structure of data documentation, by emphasizing on making data Findable, Accessible, Interoperable, and Reusable. Findability emphasizes that data and metadata should be easily discoverable. Accessibility ensures data can be retrieved using open-source web or standardized protocol. Interoperability focuses on the integration of other datasets though the use of shared format or common ontologies/vocabularies. Reusability highlights the need for data to be richly described with relevant attributes.

Adapting datasets to a purposed standard formulates back to the FAIR principle of evaluating the data of which we have been given, and to contribute a fair assessment of their complexities by providing a quality control that ranks and assesses each dataset.
For example, our assessment process of ranking data quality based on our framework was majorly facilitated by the condition of our data, nevertheless for the case of France only one set of data group was considered (INPN), a major step back as we will mention later in this paper. Our evaluation for their interpretability ranked the data of France under the value of L3 which is applicable due to their qualified metadata that makes reference to other metadata, as seen by their polygon sites that references distinct habitat attribute values that do not pertain to the original dataset. An in-depth analysis for the remaining dataset is conducted and mentioned as part of our methodological discussion found in this paper. For resulting ranking of all our data register can be seen under Appendix B of this paper.
It is important to remark that this framework only presents the data accessible at the time of processing but does not account for the effort made to access more than the information available through various data providers. Public institutions were reached out with the aim to obtain a more detail data set that aligned with our study. The Agence de l'environnement et de la maîtrise de l'énergie (ADEME) holds habitat maps containing national codes, Annex I HD, or EUNIS for the region of Languedoc-Roussillon. The request was transferred to the director of the Information and Observation Department of the Regional Committee for Tourism and Leisure of Occitanie, but there was no response. The Conservatoire d'Espaces Naturels d'Occitanie also holds habitat maps containing national codes, Annex I HD, or EUNIS for the region of Languedoc-Roussillon. The message was transferred to the INPN, but again, there was no response. The Inventaire National du Patrimoine Naturel (INPN) has data available that contains the national codes for SITES and a URL to a third webpage that relates each SITE to one or many Annex-1HD codes. A request was made to access habitat layers found on the INPN webpage, but no response was received. A personal inquiry was made to obtain a more comprehended habitat map that could contain geographical information that rage with either Annex 1 HD or EUNIS classification standards, this was unsuccessful.

Without a more detailed outlook at the data accessible required us to adapt our original methodological approach for reclassifying habitat. This research highlights the complexities and challenges in data recompilation for habitat mapping, underscoring the need for standardized methods and cooperative data sharing among institutions.

ii. Catalunya and Languedoc–Roussillon habitat map harmonization

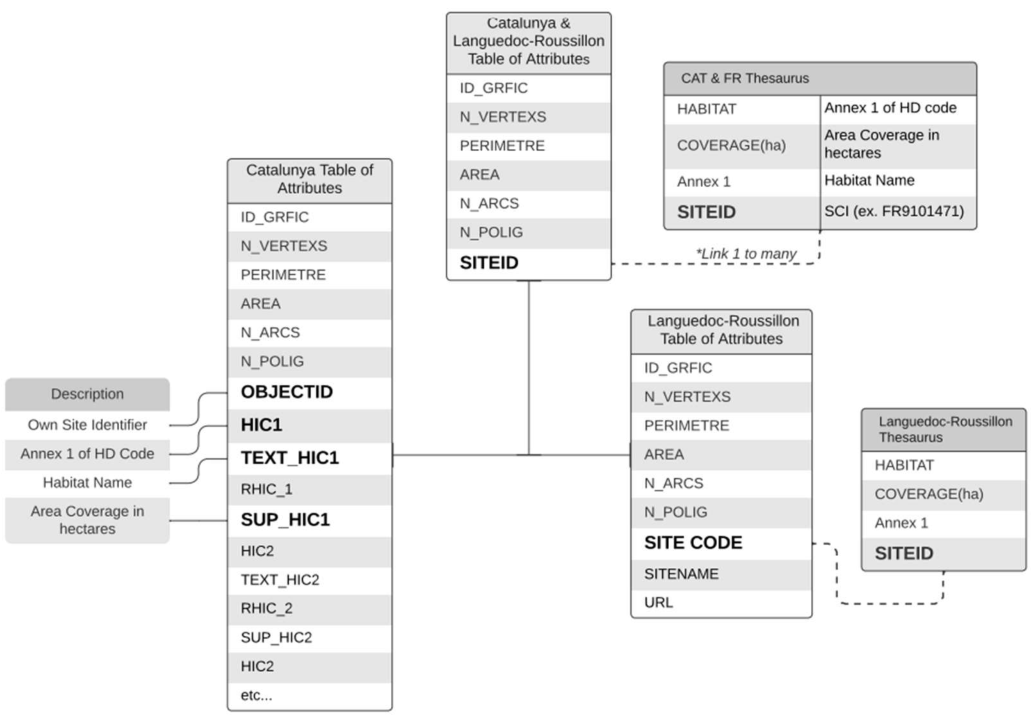


Figure 4. Scheme for unify attributes for France and Catalunya

For a harmonized map of France and Catalunya we began by analyzing our recollected public habitat data. We began by looking at the region of Catalunya, the dataset selected used to conduct this study was the "Mapping of habitats of community interest in Catalonia, version 2 (2018)" [17] which provided a well-documented file of the habitat present in that region. The table of attributes is structured as seen under Catalunya table of attributes (Figure 4), as part of our develop scheme structure for Catalunya with a OBJECTID representing a region of classification, following by a numerical code for each hàbitats d'interès comunitari (HIC) that are contained in each OBJECTID, followed by the name of distinct habitat. These habitat codes refer to Annex 1 of the European Union's Habitats Directive. RHIC & SUP_HIC values are there to attribute the covering area of the polygon into a metric of hectares (ha) and values of 0 to 10 (or 0 – 100%).

Is also important to highlight the fact that the area per polygon has been of a high quantity for Catalunya (Number of graphical objects = 23940) and also the region of Andalusia (Number of graphical objects = 95357) both of a compact size for the polygons observed (mean of ~26ha for Catalunya and ~22ha for Andalucia) (see Appendix A – Table 4 & 6), this stands out when compared to the data available for France (Number of graphical objects = 121) and even for the area of Faro, Portugal (Number of graphical objects = 3267) (see Appendix A – Table 3 & 5), which addresses quantity consistencies to be higher in Spain then other nations regions observed. For this case concluded by implementing a unique distinctive habitat reclassification approach that attempts to adapt these irregularities into a more cohesive structure.

The territory of Languedoc–Roussillon presented a challenge when it came to analyzing, as the data could not be pinpointed whether at the regional (Languedoc-Roussillon and Midi-Pyrénées) or administrative (Occitania) level. Thus, the data had to be adjusted from the national product L'Inventaire national du patrimoine naturel (INPN)[14]which provides a SIG layer containing Sites of Community Importance (SCI) and special areas of conservation (SAC). In the data available these SAC are attributed as SITECODE which is a numerical identifier assigned to a polygon (habitat site), the only thing else that contains are the name of the site, and a URL to conduct an associated search of the potential habitat enclose in that area in the INPN website. As seen by Languedoc-Roussillon table of attributes (Figured 4) even though it presents a simplified approach to habitat mapping it still faces challenges associated to the lack of metadata quality information which produces a step back for interoperability between border states products.

The French zone required a closer look to considered appropriate data to be used on this study on how to display habitats in that region. This habitats, as mentioned previously, could only be viewed though a URL website linked for each SITECODE which directed you to the INPN web page for habitats positional reference and code names, but could not be consulted via a GIS software. Many attempts were made to access these habitat layers that were available on the web through an experimental approach by looking at the XML files attached, and even personal inquiries done to the INPN as mentioned before, all were unsuccessful. This results in adaptive methodological approach of manually inputting all habitat codes and related habitat names to their designated SITECODE.

To achieve this, we began by clipping the original national layer for Languedoc–Roussillon, to decrease the amount of SITECODE's to classify by using QSIG. In the end, a total of 1293 habitats were manually encoded to be linked to all 121 SITECODE, meaning there are various habitat in one or many areas represented by the SITECODE's. This was expected seen by the fact that the mean of our polygons in terms of area projection was of 4481 ha (see Appendix – Table 5) for the region of Languedoc–Roussillon. Once we had the DBF created we proceeded to conduct a link of 1 to many as a function of the thesaurus forming part of the original table Figure 4 (Languedoc–Roussillon Table of Attributes). The result being the ability to view said habitats without consulting the web, providing a first step into unifying both tables.

To begin with the process of unification it is required for both areas to be georeferenced. For Catalunya dataset, its original coordinate reference system was UTM zone 31 Northern Hemisphere with ETRS89 datum and for Languedoc-Roussillon dataset, it was Lambert Conformal Conic with ETRS89 datum. We decided both to be georeferenced using UTM zone 31 as it encompasses much of northern east of Spain (Catalunya) and a significant portion of France, ensuring that the area of interest falls well within a single UTM Zone, and decided to change from ETRS89 datum WGS84 global datum, resulting in two layers both georeferenced in UTM-31N-WGS84 (EPSG:4326).

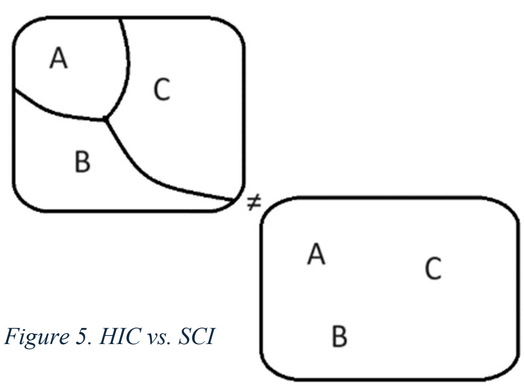


Figure 5. HIC vs. SCI

Once we had both of the maps geometrically harmonized, we began plotting methods to couple them both into a homogeneous layer in terms of habitat codes. Since France dataset pertain to a different set of table function like seen above, we were given 2 options either display a map legend with the available habitat codes or display them based on protected sites (Figure 5. HIC vs. SCI). For a more accurate and comprehensive representation, that would reduce incoherences, a priority to protected sites representation has been given based on the quality of the data available. If we presented the map to represent habitat code (instead of protected sites), the different habitats within a protected site could be listed but not spatially represented due to lack of information of the exact limits of each habitat within a site, which could be considered as an inaccuracy. The technique of mapping of site codes instead of habitat codes has been previously employed, as seen under the Nature 2000 network of Habitat Directive sites (pSCI. SCI or SAC) created by the European Environment Agency, designed to display specific attributes that indicate their conservation status and habitat protection [11].

Thus, the attributes table of Catalunya dataset had to be adjusted to reflect the desired structure for reclassification of site codes to habitat code. As seen in Catalunya table of attributes (Figure 4) table each polygon contains an OBJECTID with various habitat codes (HIC), data had to be adapted so that each habitat could reflect their position in one or many OBJECTID's. To do so we had to convert Catalunya DBF into a compatible format for processing such as Excel (CSV), to be able to manually separate their habitat values from there site position, seen that from each OBJECTID there could be one or more HIC#, and position those data set into our desired attribute table, and for this case as function of our thesaurus as our end product that contain habitat values for Catalunya and France. The conversion formats utilized the DBFCSV tool found in MiraMon software. Since the data for Catalunya contained a various amount of information and for the separation of habitat with their site codes could only be done though a manual process it also required us to narrow down our study to a rectangle of exactly 40 x 20 km long. The layer clipping was done through the use of Retalla tool in MiraMon which permitted us to determine the envelope clipping based on our exact optional coordinates indicated with MinX MaxX MinY and MaxY, which uses a metric coordinate system to operate, UTM-31N-WGS84 in this case to achieve our selected zone of 40 x 20 km.

Both DBF's were transferred into one set of table attributes that follow pattern of resulting Catalunya and Languedoc-Roussillon attribute table (Figure 4), which displays its habitat node, name and hectare size (ha). All defined by a set of OBJECTID and linked though a thesaurus that contained their habitat, unifying both maps under one set of attributes and single geographical layer.

To recap the general process of unifying distinctive attribute tables resulting in our harmonized map, consisted of our efforts to identifying 4 key template fields that could be associated for both available data: habitat code in correspondence with the Annex 1 of HD, habitat name, and size. The key factor in implementation of this methodology is the priority to include habitat coverage estimation, which relies on our principle of data reliability when comparing multi polygon sites that might enclose in one or more habitats. This allows the reader on top of being able to view the habitat name and code, easily identify the predominant value enclose in each multi polygon site for each location.

i. Faro and Andalusia habitat map harmonization

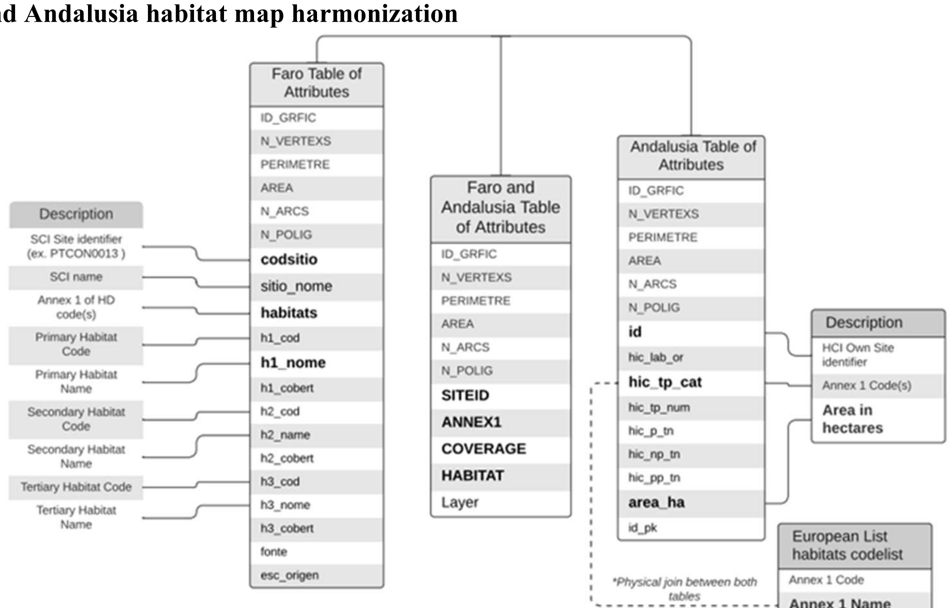


Figure 6. Scheme for unify attributes for Faro and Andalusia

For the region of Faro and Andalusia, the methodology process consisted of relatively of the same approach. Regarding the region of Faro, in the south of Portugal the data provided by the Instituto da Conservação da Natureza e das Florestas (INCF) has been used for this study, specially the Habitats da Diretiva Habitats - RN2000 (2013 - 2018)[15] product, which consisted of polygons for each "Zonas especiais de Conservação"(ZEC) accessible through a WMS. These special areas of conservation ZEC act as the SCI and HIC, contributing to the conservation of natural habitats (Annex1), by creating a coherent ecological network of Special Areas of Conservation (SACs) and Special Protection Areas (SPAs) selected on the basis of a criteria set by the Nature 2000 network. The table of attributes is constructed as seen in Faro table of attributes (Figure 6) with a "Codsitio" representing a region of classification or site-code, following by a numerical order for each habitat that are contained in each "Codsitio", followed by the name of each distinct habitat. Additionally for each polygon the viewing was in relation to each habitat that pertains to that area with their name and size.

For the region of Andalusia, the data used was the Hábitats de Interés Comunitario (HIC)[16] de Algarve, as a GPKG file originated from the Andalucian Environmental Information Network (Rediam), this platform is managed by the Junta de Andalucía Consejaría de Sostenibilidad, Medio Ambiente y Economía Azul. The table of attributes was constructed as seen under Andalusia table of attributes (Figure 6), with a "id" representing a region of classification or site-code, and not "fid" as that is an automated identification code and "id" provided the identifier for each polygon and not just an automated value. In this case "Hic_tab_cat" refers to the numerical value for each habitat that is contained in each "id", following the identifier number for Annex 1, additional important value "area_ha" which just refers to the size of the set habitat, for a better understanding of each parameters its recommended to view their metadata definition which are attached to this paper's appendix (see Appendix A - Table 6).

The unification process consists of our previous approach by beginning on georeferencing both layers. For Andalusia its original coordinate reference system was UTM zone 30 Northern Hemisphere with ETRS89 datum and for Portugal the Gauss-Krüger with ETRS89 datum. We decided both to be re-georeferenced using UTM zone 29 as it encompasses the southwest side of Spain and portion of Faro, in Portugal, resulting in a unified layer in UTM-29N-WGS84 (EPSG:32629) reference system.

A selective process of narrowing down the data begin by selecting a zone of 25 x 15 km since the amount of polygons was higher than one would have expected in the case of reclassifying and modifying attribute tables for each layer, this was a significant struggle specially when it came to analyze the Andalusian layer, as just for the region Huelva in Andalusia contain around 95357 polygons (see Appendix A - Table 6).

Finally, the approach for displacement determined by its habitat to be the best approach in comparison to habitat map created for Catalunya and France which prioritize their site codes over the legend. This was mayorly decided by the size in which the habitat were contain Andalusia had a mean of ~22 ha (see Appendix A – Table 6) and the area of Faro had a mean of ~15 ha (see Appendix A– Table 3). Both tables DBF were transformed on to a CSV to be organized in the table scheme as seen for Faro and Andalusia table of attributes (Figure 6).

Attribute table of Andalusia did not contain the name for which their habitat of their polygon pertained, and for a proper implementation of the data it became crucial to include that set of information in the overall map. To do this we had to obtain the list of Annex 1 habitats and their dictionary table that defined their value, we discover that EEA European Red List[13] of Habitats, contain an extensive dictionary table under a data sheet which required a simple modification to eliminate irrelevant data in excel such that the values of "HABITAT" and "ANNEX1" codes would be the only present (Figure 6), as this information was transferable to a DBF table set and attribute those name values to our unify table of Faro and Andalusia by performing a physical join of both tables as part of a function in MiraMon software. This automated approach of transferring values from one table set to another allows us to view not just the Annex 1 habitat codes but their official habitat names, allowing the viewer to see the habitat official label pertaining to their numerical code, as this is also an important value display for precise habitat classification. Once we adapted both data sets situated under our preferred table structure, we unified them as seen under Faro and Andalusia attribute table resulting product (Figure 6).

In summary our methodology follows our general approach of identifying and encapsulating 4 key template fields that could be associated with both available data. We focus on a design for the result map that prioritize the predominant habitat in each polygon site, as seen by the geometrical polygon site containing less than 3 habitats and seen by their compatible dimension between both tables (Faro and Andalusia). Additionally, our distinctive key factor unique to the methodology pertaining to area of Faro and Andalusia is our effort of harmonization that relies on the decision to incorporate the habitat code to their exact habitat name as to enrich the semantic understanding of complex attributes, which improves the interoperability of the data in respect of the original dataset.

4. Results

Two harmonized habitat maps have been produced: one for the "Catalunya & Languedoc-Roussillon regions" based on SCI Approach (Figure 8), and other one for the "Faro and Andalusia" regions (Figure 10).

These maps incorporate an identifier for their national habitat (as shown in the printed version in Figure 8 and Figure 9) but also contains their full habitat name, Annex 1 codes, and also a representation of their habitat in terms of hectare (ha)(see the MiraMon output MMZX maps provided attached to this paper). All under a unified coordinated reference system, scale, and legend.

To demonstrate our efforts of harmonizing habitat values and to expand upon our developed prototype we have produced an in-depth analysis that demonstrates the data enclosed for a selective group for each habitat map. Here we present a pie chart showing the habitat coverage distribution for SITEID FR9101472 (France) and 41470 (Catalunya) (see Figure 7). It illustrates the various habitat types, and their respective proportions based on the coverage area in hectares.

For the polygon site FR9101472 (France), the largest portions of the chart are dominated by habitats 9430 (Subalpine and montane Pinus uncinata forests) and 4060 (Alpine and Boreal heaths), each making up 41% (3968.8 hectares for both habitats) of the total coverage, its pie chart is complex due to the higher number of habitat types and their varying coverage sizes, is important to remark that this is not an indicator of high level of biodiversity since is focused on a large coverage area.

For the polygon 41470 (Catalunya) for this specific polygon it only contains 3 habitats being Annex 1 code 8110 (2.81 hectares) as the predominant value indicating a more balanced and narrowed down scale distribution. Both SITEID were selected based on positional nature since both are located next to the borderline that separates France and Spain. We discover that they do share similar attributes seen by containing habitat 8220 (Siliceous rocky slopes with chasmophytic vegetation), 8130 (Western Mediterranean and thermophilous scree), and 8110 (Siliceous scree of the montane to snow levels) comprising of 13 % of the polygon for France and 100 % for Catalunya (Figure 7).

We also present a closer look into the habitat distribution for Faro and Andalusia. It illustrates the various habitat types, and their respective proportions based on their visible coverage. Our area is selected based on their positional nature of being located near the borderline that separates Portugal and Spain.

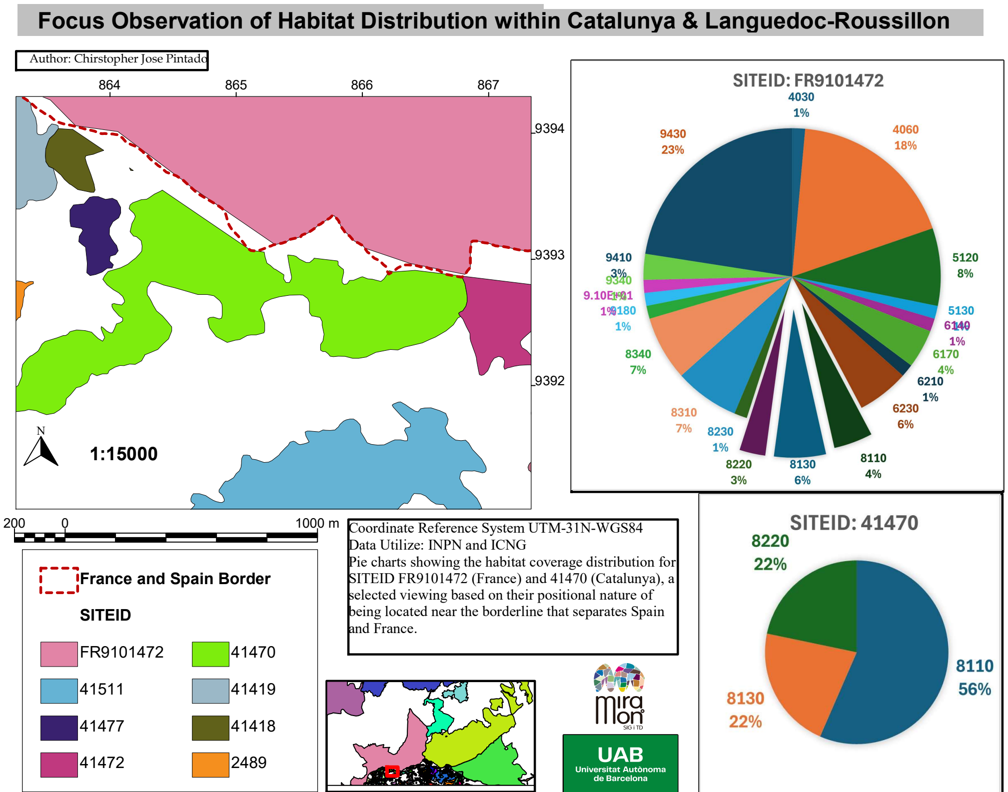
A visual analysis can shed light on the values peculiarities when observing the type of classes and their distribution when accounting borderline behavior. Our legend brings forward the habitat codes available using the Annex 1 classification, but it also addresses particularities, such as that 2 to 3 habitats are shown for a particular polygon. As explained in previous section 3, this was according to official data provided by (INCF). This approach is due to not containing an exact depiction for the habitat enclosure in their geometrical site due to the lack of precise area coverage estimation like seen in the harmonized habitat map for France and Catalunya. This results in the visualization of all the habitats to be displayed for that one specific polygon, valuing it as a unique national habitat. And lastly, we conclude for our specific focus area of Faro and Andalusia that habitat 6310 (Dehesas with evergreen Quercus spp) to be most compatible for our close-up analysis (Figure 9).

Observational sight can distinguish differences like median size for each habitat site depending on the country predominant position. What stands out after a comparison at each consolidated maps is the data originating from Spain is of a higher detail than Portugal and specially in respect to France. This is validated in Spain related habitat product (Catalunya and Andalusia) by citating in their collective metadata which adds mentions to the law 27/2006, of July 18[12] that regulates and provide the rights of access to environmental information for citizens. A comparative approach has not been made for all three nations information gathering request but is fair to promote this citation for a potential justification for the quality of the habitat information found over the Spanish region.

To recap, the resulting habitat map illustrates our two distinctive approaches of classification, France and Catalunya are design by incorporating enclose habitats under both SCI and HCI site ID's. In comparison Faro and Andalusia are developed under a unified integration model that is classified under the predominant values based on adaptable data. Both map's precise habitats can be viewed under an observative search of their attributes found under MMZX map delivered under this paper (Appendix B).

Finally, it is important to mention that attempts to provide a comparative look of various habitat classification codes besides Annex 1 were made. The trouble arose when the habitat codes (Annex 1) were not of an identical match to the counter classification EUNIS enabling a physical join of both tables, thus, to preserve the accuracy of our map relied on their values under their official classification according to their institutional authentication.

Figure 7. Focus observation of Catalunya and Languedoc-Roussillon



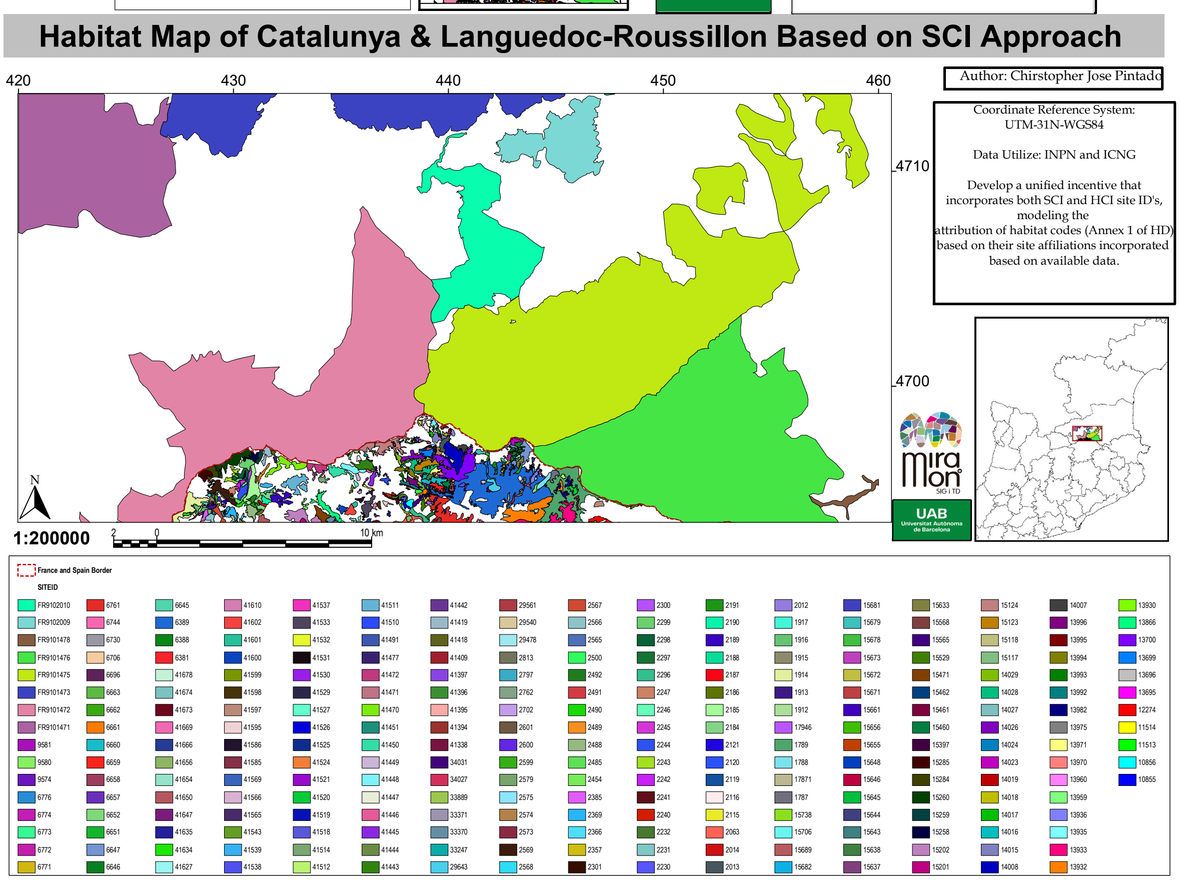
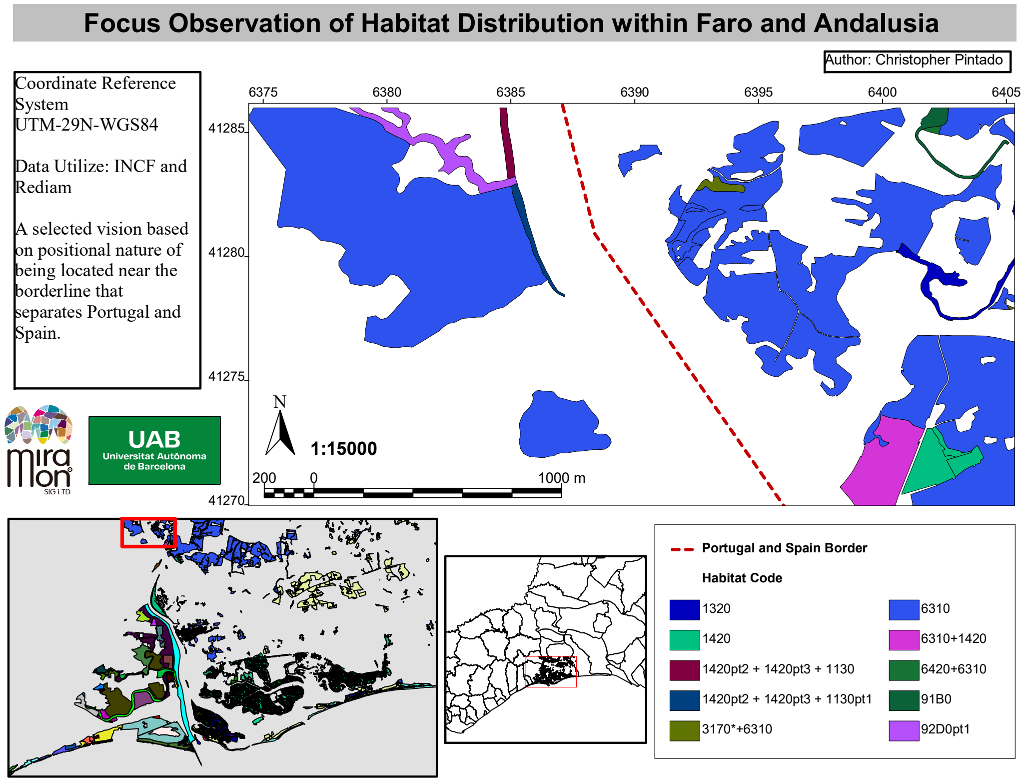


Figure 8. Habitat map of Catalunya and Languedoc-Roussillon based on SCI Approach

Figure 9. Focus Observation of Faro and Andalusia



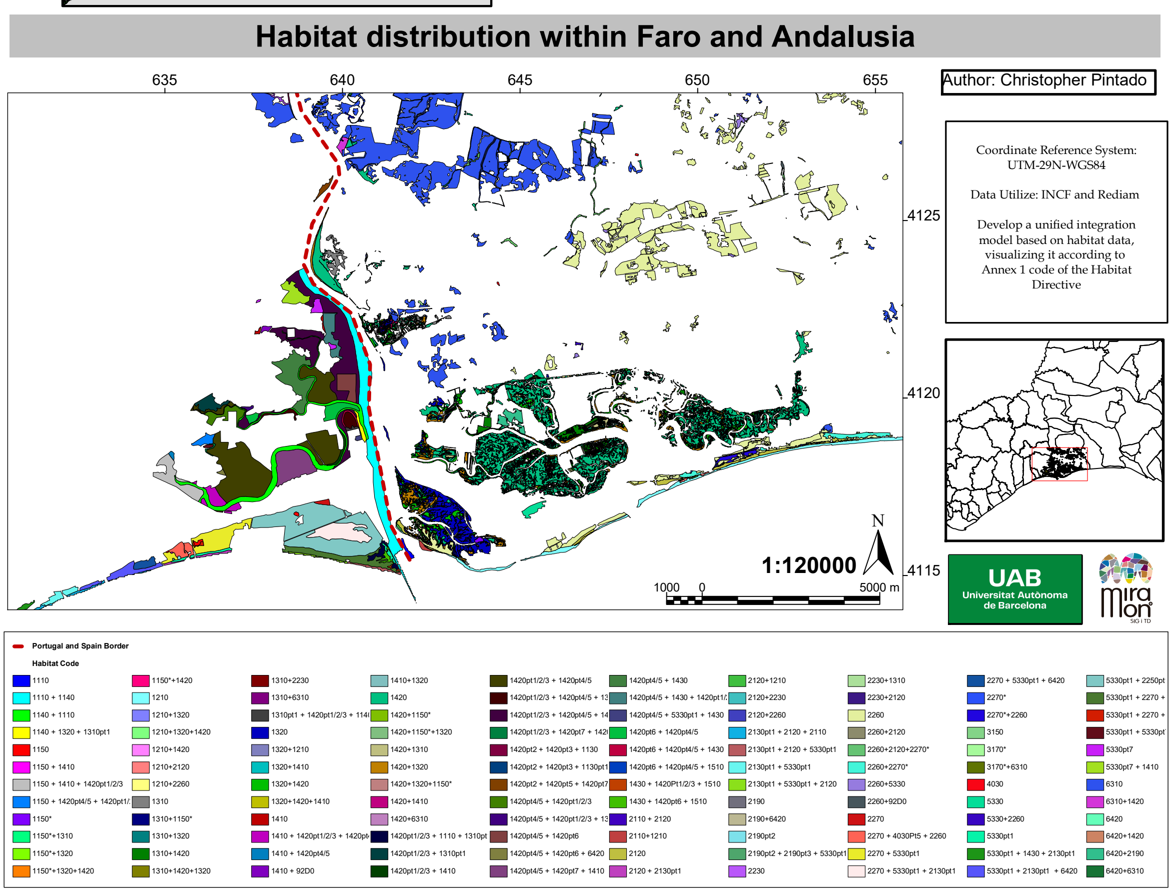


Figure 10. Habitat distribution within Faro and Andalusia

5. Conclusion

This study presents a feasible approach for harmonizing habitat maps across different European regions based on publicly accessible geographical data. By comparing and reclassifying habitat data from Faro, Andalusia-Huelva, Catalunya, and Languedoc-Roussillon regions, we have created unified maps that align with Annex 1 of the Habitats Directive codes.

Our methodology has demonstrated a significant integrational result for diverse data sets into a coherent habitat map, overcoming challenges related to linguistic differences, varied data formats, lack of meaning and institutional barriers.

Despite progress, challenges remain particularly in achieving full compatibility among different classification standards used, the study also faced difficulties in accessing detailed data from available providers. Highlighting the need for improved data sharing and cooperation among institutions.
Overall, this paper contributes to the broader goal of creating a standardized approaches for habitat mapping across Europe. This study has taken inspiration from various well stablish methodologies that have consisted of various approaches aimed to habitats harmonization as part of our scholarly review to the primal inquiry of habitat mapping at a European scale.

We presented an in-depth analysis behind all available data sets in each of the 4 regional regions. Various implementation approaches have been made to adapt said data to reflect a cross bilateral homogenize relation between European boundaries for each conterminous region, and how that approach differ to each other. We aim to provide a small glimpse of a standardized habitat classification, to provide a clue on expanding this harmonization approach to include more regions and refine the methodologies for data integration.

Future research should explore a comparative approach with general standardized habitat layers that could enrich and confirm available results. Like data from the MAES Ecosystem Type Maps created by the EEA which represents probabilities of EUNIS habitats presence in ecosystem types at a European level[20].

This TFM contributes to the InCASE project -In Situ showcase Supporting Climate and Environmental Monitoring in Europe-[1] funded by the European Commission in support to the European Environmental Agency (EEA) and the Copernicus InSitu Component. We hope this study could enhance implementation on environmental strategies and contribute to biodiversity and habitats conservation methods in Europe.

Acknowledgements

I would like to express my sincere gratitude to my thesis advisor, Joan Masó, for his invaluable feedback and expertise, and to Alba Brobia for your insightful guidance and encouragement throughout this research, this TFM would not be possible without them.

I also extend my thanks to the professors and coordinators of this master's for their dedication and for providing such an enlightening experience.

And to my family whose love has been and continues to be my greatest source of strength.

References

[1]. *European Commission, Directorate-General for Environment, (2021). EU biodiversity strategy for 2030: bringing nature back into our lives, Publications Office of the European Union. https://data.europa.eu/doi/10.2779/677548*

[2]. *European Council. (1992). Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. Official Journal of the European Communities, L 206, 7–50. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31992L0043*

[3]. *European Environment Agency. (2020). EUNIS habitat classification. Retrieved from https://eunis.eea.europa.eu/habitats.jsp*

[4]. *QGIS Development Team. (2024). QGIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved from https://qgis.org*

[5]. *Pons, X. (1994). MiraMon: Sistema d'Informació Geogràfica i software de Teledetecció. SoftwarePons. V 9.1a*

[6]. *Comparison of Land Cover/Land Use and Habitat Classification Systems for Habitat Mapping from Space: Strengths and Weaknesses Evidenced in Mediterranean Sites of Natura 2000 Network. (2013). Tomaselli Et Al. https://doi.org/10.1553/giscience2013s560*

[7]. *Volpato, A., Buckley, C., & Moran, J. (2024). Assessing and mapping habitat quantity and quality in High Nature Value (HNV) agricultural landscapes. Journal for Nature Conservation, 78, 126568. https://doi.org/10.1016/j.jnc.2024.126568*

[8]. *Adamo, M., Tarantino, C., Tomaselli, V., Veronico, G., Nagendra, H., & Blonda, P. (2016). Habitat mapping of coastal wetlands using expert knowledge and Earth observation data. Journal of Applied Ecology, 53(5), 1521–1532. https://doi.org/10.1111/1365-2664.12695*

[9]. *Nicolau, R., Basos, N., Marcelino, F., Caetano, M., & Pereira, J. M. C. (2020). Harmonization of categorical maps by alignment processes and thematic consistency analysis. AIMS Geosciences, 6(4), 473–490. https://doi.org/10.3934/geosci.2020026*

[10]. *Slootweg, J., Hettelingh, J., Tamis, W., Van 't Zelfde, M., & Posch, M. (2005). Harmonizing European land cover maps. Netherlands Environmental Assessment Agency, 47–70. https://scholarlypublications.universiteitleiden.nl/access/item%3A2882641/view*

[11]. *European Environment Agency. (2023, October 25). Protected sites in Europe (Natura 2000).https://www.arcgis.com/home/item.html?id=fa8b41fdc9e74841a054a4946e944717*

[12]. *Jefatura del Estado. (2003, July 18). Ley 27/2006. Referencia: BOE-A-2006-13010. https://www.boe.es/eli/es/l/2006/07/18/27/con*

[13]. *EEA. (2022, May). European Red List of Habitats. Prod-ID: DAT-160-en. https://www.eea.europa.eu/en/datahub/datahubitem-view/de2276d8-e295-4cd7-89c9-88812065db87?activeAccordion=*

[14]. *Inventaire National du Patrimoine Naturel(INPN). (2024). Nature 2000 Data Search. https://inpn.mnhn.fr/accueil/recherche-de-donnees/natura2000*

[15]. *Instituto da Conservação da Natureza e das Florestas (INCF). (2024). Directive Habitat. https://www.icnf.pt/conservacao/redenatura2000/diretivahabitats*

[16]. *Junta de Andalusía. (2023). Información de Referencia de los Hábitats de Interés Comunitario terrestres (IRHICA). https://www.juntadeandalucia.es/medioambiente/portal/areas-tematicas/biodiversidad-y-vegetacion/habitats/informacion-referencia-habitats-interes-comunitario-terrestres-irhica*

[17]. *Institut Cartogràfic i Geològic de Catalunya (ICGC). (2018). Cartography of habitats of community interest in Catalonia. https://catalegs.ide.cat/geonetwork/srv/api/records/habitats-interes-comunitari-v2-2018?language=eng*

[18]. *Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3. https://doi.org/10.1038/sdata.2016.18*

[19]. *European Parliament. (2003, May 26). Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS). http://data.europa.eu/eli/reg/2003/1059/oj*

[20]. *EEA. (2015, December). Ecosystem type map V2.1 (EUNIS Level 2. https://bio.discomap.eea.europa.eu/arcgis/rest/services/MAES/MAES_ecosystem_type_maps_WM/MapServer*

[21]. *Brobia, A., Masó, J. (2024, March 15). InCASE: In Situ Showcases Supporting Climate and Environmental Monitoring in Europe. InCASE-D3.2 SC3. CREAF. EEA*

Appendix A: Additional statistics and metadata details

Table 3. Faro Statistics

Àrea del polígon (projecció): Faro_HZEC_WGS84P.dbf
N records with data: 3267
Mode: 3.4255 m² (0.0003 ha) (mode is not unique)
Mean: 151393.4031 m² (15.1393 ha)
St. devi.: 340882.3681 m² (34.0882 ha)
Variance: 116200788902.5655
Summation: 494602247.8337 m² (49460.2248 ha)
Minimum: 3.4255 m² (0.0003 ha)
Maximum: 6477688.9197 m² (647.7689 ha)
Range: 6477685.4941 m² (647.7685 ha)
Median: 58320.2067 m² (5.8320 ha)
Mean absolute deviation around the median: 126412.3026 m² (12.6412 ha)
N unique values: 3267

Table 4.Catalunya Statistics

Area of the polygon (projection):BCN_HIC_WGS83_2P.dbf
N records with data: 23940
Mode: 0.6897 m² (0.0001 ha) (mode is not unique)
Mean: 266775.3129 m² (26.6775 ha)
St. devi.: 942993.2084 m² (94.2993 ha)
Variance: 889236191161.9644
Summation: 6386600990.7466 m² (638660.0991 ha)
Minimum: 0.6897 m² (0.0001 ha)
Maximum: 53005003.6074 m² (5300.5004 ha)
Range: 53005002.9177 m² (5300.5003 ha)
Median: 93550.7667 m² (9.3551 ha)
Mean absolute deviation around the median: 218318.6866 m² (21.8319 ha)
N unique values: 23940

Table 5: Languedoc–Roussillon Statistics

Área del polígono (proyección) (N) (D:\TFM_Research\STAGE2\FRANCE\sic_ue\Miramon\Languedoc-RoussillonP.dbf):
N records with data: 121
Mode: -0.0 m² (-0.0 ha) (mode is not unique)
Mean: 44812579.3 m² (4481.3 ha)
St. devi.: 73516543.8 m² (7351.7 ha)
Variance: 5404682217866792.0
Summation: 5422322090.8 m² (542232.2 ha)
Minimum: -0.0 m² (-0.0 ha)
Maximum: 396430716.6 m² (39643.1 ha)
Range: 396430716.6 m² (39643.1 ha)
Median: 5786617.9 m² (578.7 ha)
Mean absolute deviation around the median: 43595745.6 m² (4359.6 ha)
N unique values: 121

Table 6: Huelva Andalusia, statistics

Area of the polygon (projection): Huelva_HIC_WGS840000P.dbf
N records with data: 95357
Mode: 209.8738 m² (0.0210 ha)
Mean: 47442.6857 m² (4.7443 ha)
St. devi.: 221720.0130 m² (22.1720 ha)
Variance: 49159764171.3342
Summation: 4523992175.9757 m² (452399.2176 ha)
Minimum: 0.1418 m² (0.0000 ha)
Maximum: 19876092.9189 m² (1987.6093 ha)
Range: 19876092.7771 m² (1987.6093 ha)
Median: 5768.6313 m² (0.5769 ha)
Mean absolute deviation around the median: 45768.3662 m² (4.5768 ha)
N unique values: 95245

Table 7: HIC Andalucía table definitions

NOMBRE FI'SICO	hic_cápá_unicá	
NOMBRE	ALIAS	DESCRIPCIÓN
fid	Identificador único automático del registro	Identificador del registro por defecto
id	Identificador único del registro	Identificador asignado para cada polígono, de referencia
hic_tp_pct	Hábitats Interés Comunitario_label (etiqueta)_origen	Etiqueta HIC origen obtenida de la label(etiqueta) SIPNA origen, se presentan los Hics con su subtipo correspondiente.
hic_tp_pct	Hábitats Interés Comunitario_tipos (presentes)_European Union_porcentaje (de ocupación)	Tipos de HIC presentes en el polígono según los listados de la Unión Europea
hic_tp_cat	Hábitats de Interés Comunitario_tipos(pr ese ntes)_European Union_categoría	Tipos de HIC presentes en el polígono según los listados de la UE con su categoría de protección tras el código.
hic_tp_num	Hábitats Interés Comunitario_tipos(p rese ntes)_numero	Número de tipos de HIC presentes en el polígono según listados UE (sin subtipos)
hic_p_tn	Hábitats InterésComunitario_ prioritarios_ tipos (presentes) numero	Número de tipos de HIC PRIORITARIOS presentes en el polígono según listados UE (sinsubtipos)
hic_np_tn	Hábitats Interés Comunitario_no prioritarios_tipos(pre sent es)_numero	Número de tipos de HIC NO PRIORITARIOS presentes en el polígono según listados UE (sin subtipos)
áreá_há	Área (del polígono)_hectáreas	

Appendix B: Data Documentation and MMZX maps

[TFM_Data\u0026Maps.zip](TFM_Data\u0026Maps.zip)

ID	Member country	Geographic scope	Geographic extent	Dataset name english *Dataset name original*	Group	Short description / Resumé	Data provider	Reference year(s) available	Geometry	Format	CRS/EPSG	Spatial Resolution/Scale	Thematic coverage	Time series available	Time series steps	Legal constraints	License	URL - please specify whether: 'URL page, 'Metadata URL' or 'Data URL'	Reference /Technical documentation	Brief description by data provider	Contributor	Type of habitats classification used	potential correspondence to EUNIS Level	Find-ability	Access-ibility	Interoper-ability	Reuse
1	Portugal	National	Portugal	Habitats Directive Report (Natural Habitats)	Flora & Fauna	Natura 2000 Netw	sig-icnf	2021	Polygon	SHP	EPSG:37	1;5000	National	no	unknow	public	N/A	Data:https://si.icnf.pt/geoserverplinia/BDG/wfs?VERSION=2.0.0 Metadata: https://www.gbif.org/dataset/dff56954-d92b-4d83-b165-17c07e233b79 Metadata: https://snig.dgterritorio.gov.pt/rndg/srv/api/records/b90874cd-3d50-423c-b8d2-1264ea87ef36/formatters/snig-view		Distribution of habitats protected by the Habitats Directive for mainland Portugal. Information obtained through work carried out within the scope of the National Report on the Implementation of the Habitats Directive (2007-2012) and (2013-2018)	EEA	national code, EUNIS, Annex 1, CORINE	Level 2	F4	A1.1	L1	R1.1
2	Portugal	Regional	Portugal	Habitats das Zonas Especiais de Conservação Algarve	Protected Areas	Detail for the SACs of the mainland	sig-icnf	2023	Polygon	SHP`	EPSG:37	1;100000	National	no	unknow	public	N/A	Data; https://sig.icnf.pt/portal/home/item.html?id=3aefff2169744f9184e9f951cd2a4f2b Metadata; https://sig.icnf.pt/portal/sharing/rest/content/items/3aefff2169744f9184e9f951cd2a4f2b/info/metadata/metadata.xml?format=default&output=html		Distribution of natural habitats protected (Annex 1) by the Habitats Directive for Portugal. Information obtained through work carried out within the scope of the National Report on Implementation of the Habitats Directive.	EEA	Annex1	Level 2	F4	A1.1	L1	R1.1
3	Spain	Local	Cataluña	Habitats of community interest (CHIC)	Habitats	The legend is the list of habitats in Annex I of the European Union Directive 97/62/EC.	Gencat	2018	Polygon	GML, SH	EPSG: 25	834:20:00		no	unknow	public	N/A	URL & Data: https://agricultura.gencat.cat/ca/serveis/cartografia-sig/bases-cartografiques/habitats/habitats-catalunya/ Metadata: https://sig.gencat.cat/metadades/geonetworkMetadataDetailUuid.html?uuid=TMP0000A_27913			EEA	Annex 1	Level 2	F2	A1	L1	R1.1
4	Spain	Local	Andaluc	Habitats of Community Interest (CHIC)	Habitats	Shows a list of the Habitat of Community Interest present in each tile presented for Andalusia.	REDIAM	2021	Polygon	WMS, SH	*EPSG:3042*			no	unknow	publlic	N/A	Data and Metadata: https://portalrediam.cica.es/geonetwork/static/api/records/442ac9f3-c7c5-429a-8437-4a8422ef38f8		Rediam (Red de Información Ambiental de Andalucía) se encarga de integrar y difundir toda la información generada por los diferentes centros productores, tanto públicos como privados.	EEA	Annex 1	Level 2	F2	A1	L1	R1.1
5	Spain	Local	Barcelor	Terrestrial habitats of the province of Barcelona	Terrestial Habitats	This is a map promoted and coordinated by the Diputació de Barcelona and executed by the University of Barcelona between the years 2019 and 2022	UB	2019-202	Polygon, Single Layer, and Points	GeoPakage, SHp	UTM-31N-ETRS89	1:25,000	no	no	unknow	punlic	Creative Commons - Attribution (by)	https://dadesobertes.diba.cat/datasets/habitats-terrestres-de-la-provincia-de-barcelona		coordinated by the Diputació de Barcelona and executed by the University of Barcelona	ICGC	Annex 1, EUNIS, HIC, CORINE	Level 3	F2	A1	L1	R1.1
6	France	National	France	Sites classified under the Habitats Directive: perimeters transmitted to the EC (ZSC/pSIC/SIC)	Zones of Importance	Zones of Importance with Code classification (URL) which can be examinate under the INPN website	INPN	2023	Raster	SHP, WMS, WFS	EPSG: 2154		France, s	2001 - 2023		Punlic	N/A	Data and Source: https://inpn-inspire.mnhn.fr/catalogue/srv/fre/catalog.search#/metadata/651fcde4-5554-4777-a828-fdbddc610478		directive, proposals for sites of community importance (pSIC) transmitted by France to the European Commission are subject to publication as sites of community importance (SIC) in the official journal of the European Union. The SICs must then ultimately be designated as special areas of conservation (ZSC) by ministerial decree and publication in	DREAL	Annex 1	N/A	F4	A1.1	L3	R1.2